# Online Appendix to "Estimating the Fed's Unconventional Policy Shocks"

Marek Jarociński*

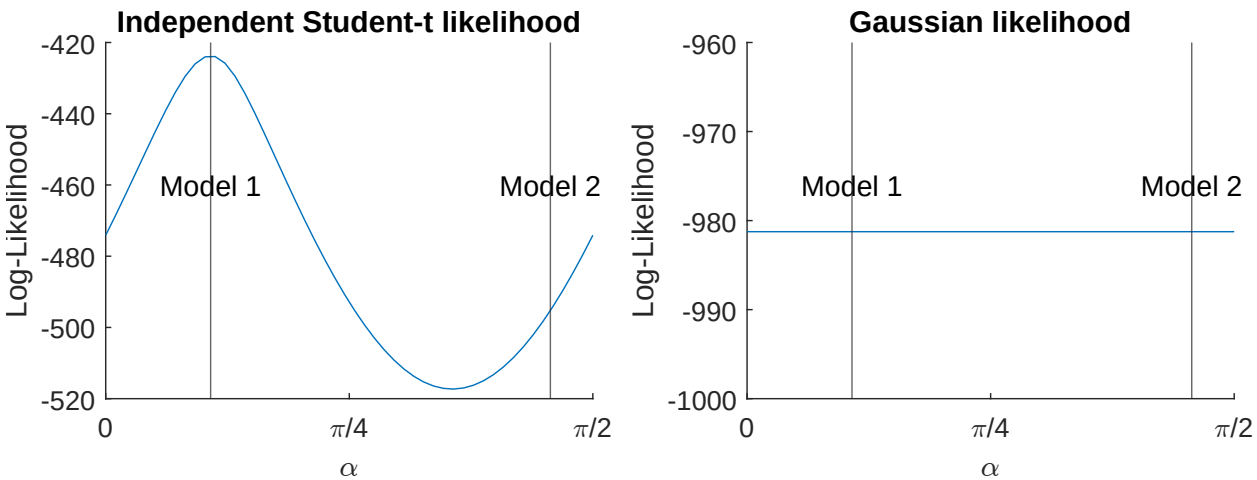
January 4, 2024

This Online Appendix contains supplementary material for the paper "Estimating the Fed's Unconventional Policy Shocks" appearing in the Journal of Monetary Economics.

*European Central Bank DG-Research, email: marek.jarocinski@ecb.europa.eu.

# A Informativeness of the likelihood function in the stylized example

Figure A.1: Stylized example: information in the likelihood function of the data from panel E of Figure 2.




Note. Likelihood of the sample in panel E of Figure 2, as a function of the rotation angle $\alpha$. $\alpha = 0$ corresponds to the Choleski decomposition of the sample variance of $Y$. Left panel: Independent Student-t likelihood given in (5). Right panel: Gaussian likelihood.

Let $Y$ be the $T \times 2$ matrix collecting the data on prices and quantities from panel E of Figure 2, i.e. generated from Model 1. Let $U$ be the $T \times 2$ matrix with orthogonal shocks. We can decompose $Y$ into orthogonal shocks in infinitely many ways because

$$Y = UC = UQ(\alpha)Q(\alpha)'C = \tilde{U}\tilde{C} \qquad \text{for any} \qquad Q(\alpha) = \begin{pmatrix} \sin\alpha & \cos\alpha \\ -\cos\alpha & \sin\alpha \end{pmatrix} \tag{A.1}$$

where $U'U = I = \tilde{U}'\tilde{U}$. Parameter $\alpha$ indexes all models that fit the data $Y$ while implying different slopes of demand and supply. All these models have the same likelihood if we incorrectly assume that the shocks are Gaussian. However, the Student-t likelihood discriminates between these alternative models. This is illustrated in Figure A.1. The log-likelihood of the data from panel E implied by the Student-t distribution of shocks, given in (5), peaks at the rotation angle $\alpha$ that corresponds to Model 1. On the other hand, a researcher who wrongly assumes the Gaussian model would not be able to discriminate between the models, as the Gaussian likelihood is the same for any value of $\alpha$. The Gaussian likelihood depends only on the first two moments and all values of $\alpha$ yield the same first two moments. However, incorrect values of $\alpha$ imply that demand and supply shocks must exhibit particular relations (such as a positive co-kurtosis), in order to match the data in panel E. This violates the independence of the shocks and hence gets penalized in the Independent Student-t likelihood.

# B Computational details for the baseline model

## B.1 The gradient of the likelihood function in the baseline model

I derive the analytical gradient of the log-likelihood (5) with the help of the results in Magnus and Neudecker (2019) and Khatri and Rao (1968). Differentiating (5) w.r.t. $\operatorname{vec} W$ yields

$$\frac{d \log p(Y|W,v)}{d \operatorname{vec} W} = T \operatorname{vec} W^{-1\prime} + \iota_T'(A \bullet Y) \tag{B.1}$$

where $\bullet$ denotes the row-wise Khatri-Rao product,

$$A \bullet Y = \begin{pmatrix} a_1' \otimes y_1' \\ \dots \\ a_T \otimes y_T \end{pmatrix}, \tag{B.2}$$

$a_t$ is an $N \times 1$ vector with the $n$-th element

$$a_{t,n} \equiv -\frac{v_n + 1}{v_n} \frac{u_{t,n}}{1 + u_{t,n}^2/v_n}, \tag{B.3}$$

and $\iota_T$ denotes a $T \times 1$ vector with each element equal to 1.

Differentiating (5) w.r.t. $v_n$ yields

$$\frac{d \log p(Y|W,v)}{dv_n} = -\frac{1}{2}\sum_{t=1}^{T} \log(1 + u_{t,n}^2/v_n) + \frac{v_n+1}{2v_n^2}\sum_{t=1}^{T} \frac{u_{t,n}^2}{(1 + u_{t,n}^2/v_n)} + T\frac{d \log c(v_n)}{dv_n} \tag{B.4}$$

where

$$\frac{d \log c(v_n)}{dv_n} = -\frac{1}{2v_n} - \frac{1}{2}\psi(\tfrac{v_n}{2}) + \frac{1}{2}\psi(\tfrac{v_n+1}{2}) \tag{B.5}$$

where $\psi$ denotes the digamma function (i.e. the derivative of the log of the Gamma function).

In practice, I reparameterize the log-likelihood in terms of $z_n = log(v_n)$.

## B.2 Bayesian estimation

For the Bayesian estimation I specify priors for parameters $W, v$. The prior for $W$ is flat, $p(W) \propto 1$. The priors for $v_n$ need to be sufficiently informative to ensure the propriety of the posterior (see Bauwens and Lubrano (1998) for a detailed discussion). I use Gamma priors with a high mean to ensure that the priors are conservative, in the sense that, if anything, they push $v_n$'s towards relatively large values where the density is closer to Gaussian and identification is weaker. This ensures that any identification comes from the data and not from the prior. The Gamma density is given by $\mathcal{G}(\alpha_n, \beta_n) = \Gamma(\alpha_n)^{-1}\beta_n^{-\alpha_n} v_n^{\alpha_n - 1}\exp(-v_n/\beta_n)$, and has mean $\alpha_n\beta_n$ and variance $\alpha_n\beta_n^2$.

I find a posterior mode $\tilde{\theta} = (\operatorname{vec}\tilde{W}', \tilde{v})'$ and the Hessian at the mode $\tilde{\mathcal{H}}$. I start the simulation from the posterior mode $\tilde{\theta}$. I generate proposal draws with a Gaussian random walk model with the innovation variance equal to $\tilde{\mathcal{H}}^{-1}$ scaled to ensure the acceptance rate of about 20%. The scale is 0.86 in the baseline model. I generate 10,000,000 draws and keep every 5,000-th. This simulation takes less than 10 minutes on a standard laptop. The convergence of the Markov Chain is confirmed with the Geweke (1992) diagnostics.

# C Normalization: further details and examples

## C.1 Further details

Let $\tilde{W}$ denote a draw of $W$, let $q = 1, ..., 2^N N!$ index the signs and permutations of the $N$ columns of $\tilde{W}$, let $\tilde{W}_q$ denote the matrix obtained by the $q$-th sign swap and permutation of the columns of $\tilde{W}$, let $\mathcal{V}_W$ denote the asymptotic variance of $\operatorname{vec} W$ (i.e., the corresponding $N^2 \times N^2$ block of the inverse of the Hessian of the likelihood at $\hat{W}, \hat{v}$) and let $F(x|m, V)$ denote the multivariate Gaussian density with mean $m$ and variance $V$ evaluated at the point $x$.

**Algorithm 1** *Given a draw $\tilde{W}$, for each signs and permutation $\tilde{W}_q$, $q = 1, ..., 2^N N!$ evaluate $f(q) = F(\operatorname{vec} \tilde{W}_q | \operatorname{vec} \hat{W}, \mathcal{V}_W)$. Take the $\tilde{W}_{q*}$ where $q* = \arg\max_q f(q)$ as the normalized $\tilde{W}$.*

Algorithm 1 can be accelerated by using the sign normalization of Waggoner and Zha (2003). While their procedure is designed for Gaussian Structural VARs, it works also with the Gaussian approximation of the likelihood proposed here. The following Algorithm 2 is faster and in the example below it produces for each draw the same normalization as Algorithm 1. Let $p = 1, ..., N!$ index the permutations of the $N$ columns of $\tilde{W}$, let $\tilde{W}_p$ denote the matrix obtained by the $p$-th permutation of the columns of $\tilde{W}$.

**Algorithm 2** *Given a draw $\tilde{W}$, for each permutation $\tilde{W}_p$, $p = 1, ..., N!$:*

1. *Scale the columns of $\tilde{W}_p$ by +/-1 using the Likelihood Preserving normalization of Waggoner and Zha (2003) (their Algorithm 1), obtaining a sign-normalized matrix $\tilde{W}_p^{LP}$.*

2. *Evaluate $f(p) = F(\operatorname{vec} \tilde{W}_p^{LP} | \operatorname{vec} \hat{W}, \mathcal{V}_W)$.*

*Take the $\tilde{W}_{p*}^{LP}$ where $p* = \arg\max_p f(p)$ as the normalized $\tilde{W}$.*

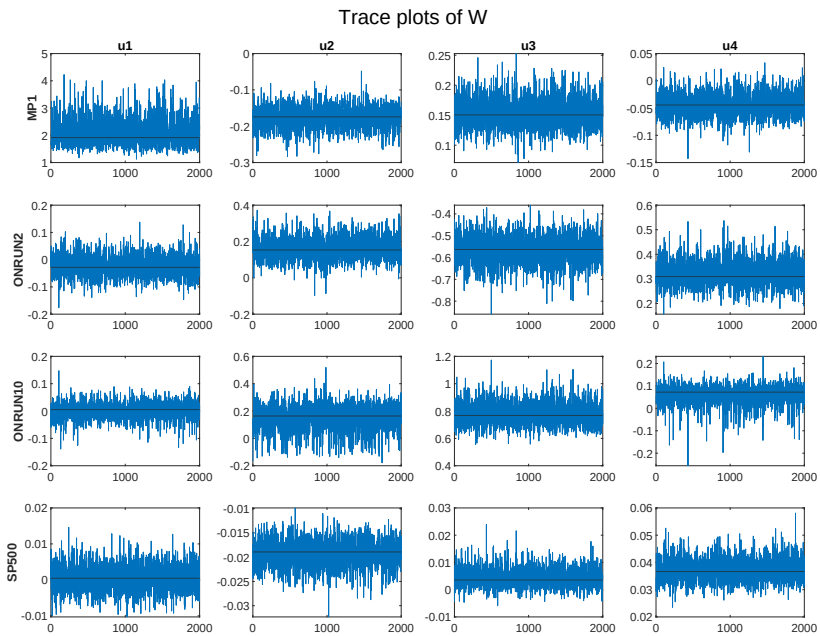
In practice, a finite Markov Chain may visit the neighborhoods of only a subset of modes. In this case, rather than considering all the $N!$ permutations, one only needs to consider the signs and permutations of those columns of $W$ that have multiple modes before the normalization. This allows to speed up the normalization in case the model is large.

## C.2 Examples

It turns out that in most of the estimations reported in the paper and in its appendices the modes are well separated and posterior simulators with standard settings never jump to the neighborhood of another mode. For example, Figure C.1 reports the trace plots of $W$ from the baseline estimation. All draws correspond to the same signs and permutation and do not require normalization.

One of the few exceptions is the estimation of the baseline model on the subsample 1991-2004, reported in the left panel of Figure F.1. (By contrast, the estimation on the 2005-2019 subsample, reported in the right panel of the same figure, does not require normalization.) Figure C.2 reports trace plots of $W$ estimated on the 1991-2004 subsample. The top panel

Figure C.1: Trace plot of $W$, estimation on the full sample (241 obs.): no normalization needed



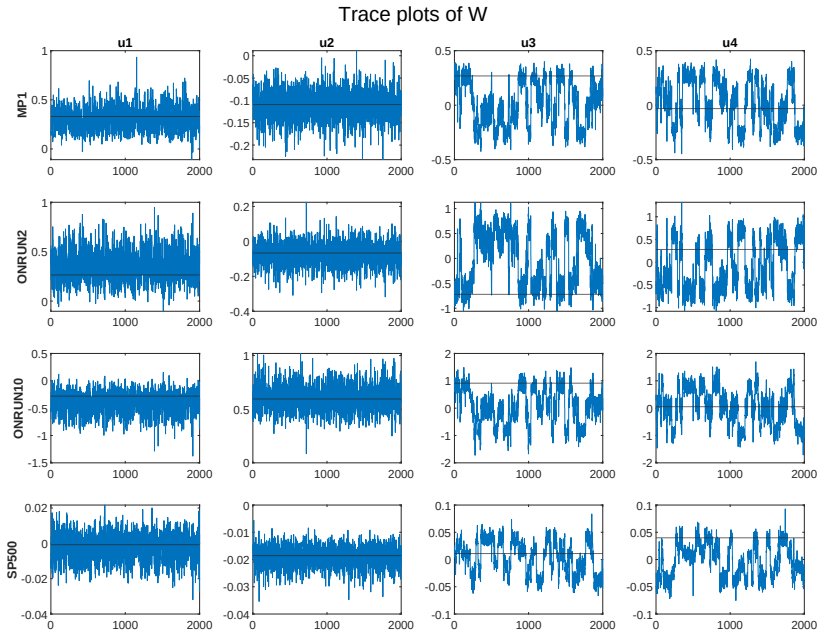

Note. The horizontal lines in the plots indicate the maximum likelihood estimates.

shows the trace plots before normalization. The trace plots of the coefficients corresponding to the shocks $u_1$ and $u_2$ show no anomalies. However, the trace plots for $u_3$ and $u_4$ fluctuate between different values, corresponding to opposite signs of the shocks or their alternative ordering. The bottom panel shows the trace plot obtained after applying Algorithm 1 or, equivalently, Algorithm 2 to each draw. Here also for $u_3$ and $u_4$ most draws cluster in the vicinity of the maximum likelihood estimates.

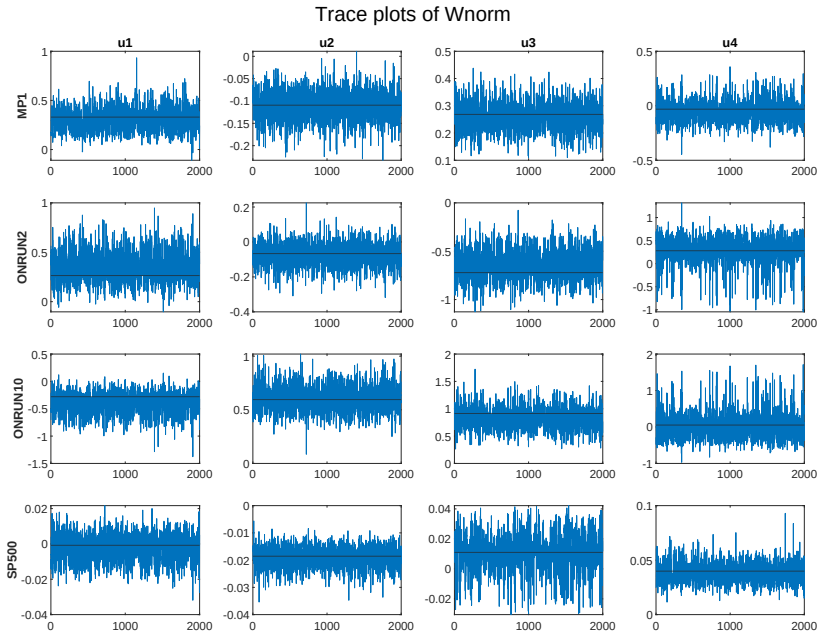
Another example is the estimation with PDMT shocks and the restriction $\bar{v} = 0.4$, discussed in G and presented in the middle panel of Figure G.2.

Figure C.2: Trace plot of $W$, estimation on the subsample 1991-2004 (120 obs.)

Before normalization

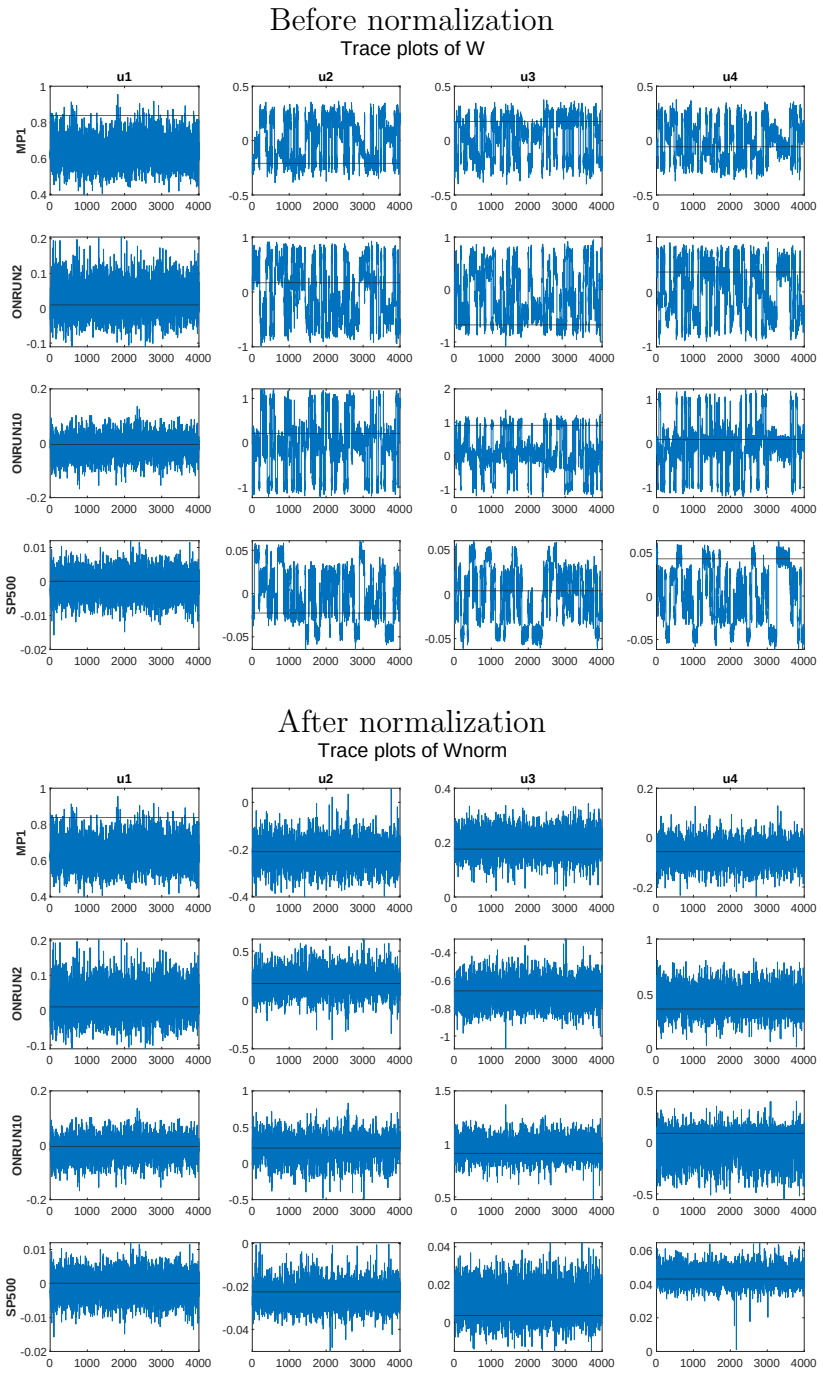



After normalization




Note. The horizontal lines in the plots indicate the maximum likelihood estimates.

Figure C.3: Trace plot of $W$, estimation with PDMT shocks




Note. The horizontal lines in the plots indicate the maximum likelihood estimates of the baseline model.

# D Baseline model: estimation and further results

## D.1 Maximum likelihood and Bayesian estimation

I maximize the likelihood using 100 random starting points and all maximizations converge to the same point, up to sign and order swap (and a small numerical error). To obtain one starting point I start with the choleski factor of the covariance of $Y$ and rotate it with a random $N \times N$ orthogonal matrix drawn from the Haar measure (Rubio-Ramírez et al., 2010). I use the inverse of the resulting matrix as the starting point for $W$. I draw the initial $v$ from the uniform distribution on the interval [1,30]. The resulting 100 modes correspond to almost as many sign and order permutations, so I normalize them to the same sign and order before comparing them. After normalization they turn out to be the essentially same. The rank correlations between the resulting shocks and the reference shocks never fall below 0.9999. The maximum difference between the respective $v_n$'s is 0.013. These are only numerical differences and there is no evidence of any nontrivially different local modes or other problems with the identification.

To cross-check the maximum likelihood estimation and to construct precise error bands I obtain 2000 approximately uncorrelated draws from the Bayesian posterior. For the Bayesian estimation I specify conservative priors for $v_n$, with $\alpha_n = 2$ and $\beta_n = 5$ for all $n$, implying the mean of 10 (which is much higher than the point estimates $\hat{v}$). For $v_1$ some draws in the Monte Carlo simulation are small causing numerical problems, so to prevent this I truncate the prior to $v_n > 0.45$. The stored 2000 draws display no significant autocorrelation and the chain is stationary according to the Geweke (1992) diagnostics. The simulation takes less than 10 minutes on a standard laptop.

## D.2 Further results

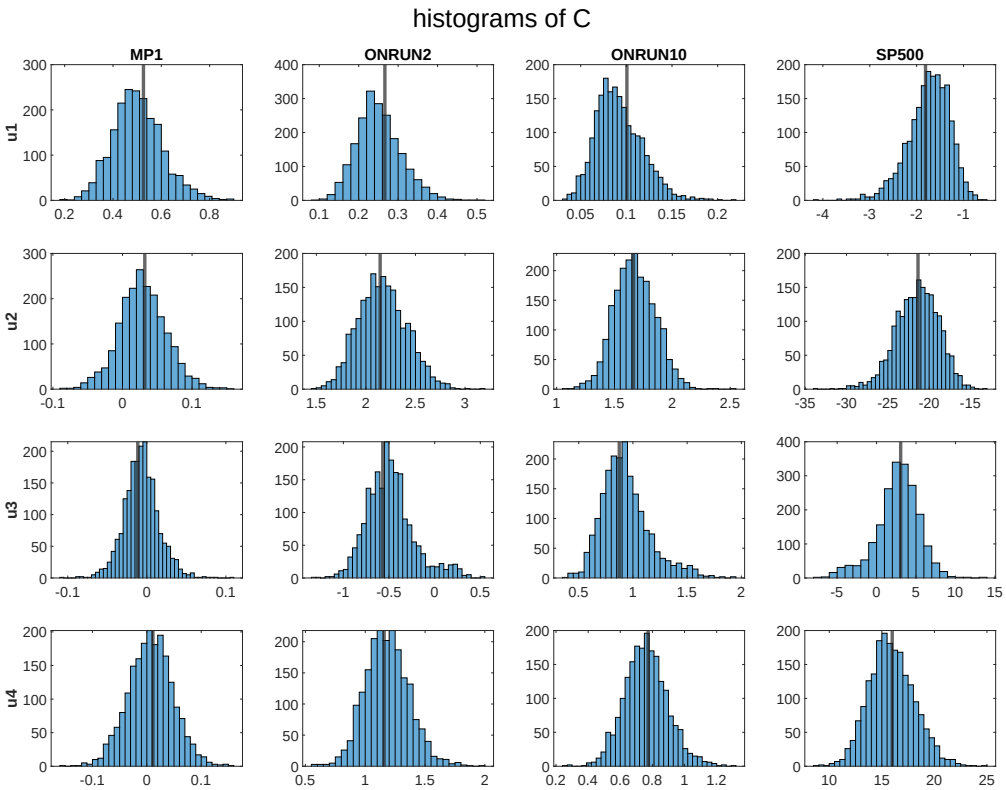
Table D.1: Weight matrix $W$

| | $u_1$ | $u_2$ | $u_3$ | $u_4$ |
|---|---|---|---|---|
| MP1 | 1.92 | -0.17 | 0.15 | -0.04 |
| | (0.46) | (0.03) | (0.02) | (0.02) |
| ONRUN2 | -0.03 | 0.15 | -0.56 | 0.31 |
| | (0.04) | (0.06) | (0.06) | (0.05) |
| ONRUN10 | 0.01 | 0.16 | 0.77 | 0.07 |
| | (0.03) | (0.09) | (0.08) | (0.04) |
| SP500 | 0.00 | -0.02 | 0.00 | 0.04 |
| | (0.00) | (0.00) | (0.00) | (0.00) |

Note: Maximum likelihood estimates. Posterior standard deviations in parentheses.

Table D.1 reports the matrix $W$, which determines how the variables are weighted to obtain shocks (recall that $U = YW$). We can see that $u_1$ is essentially equal to MP1 - it only places weight on MP1, while the weights on the remaining variables are negligible. By contrast, the other shocks are constructed from multiple variables.

Figure D.1: Posterior distribution of $C$



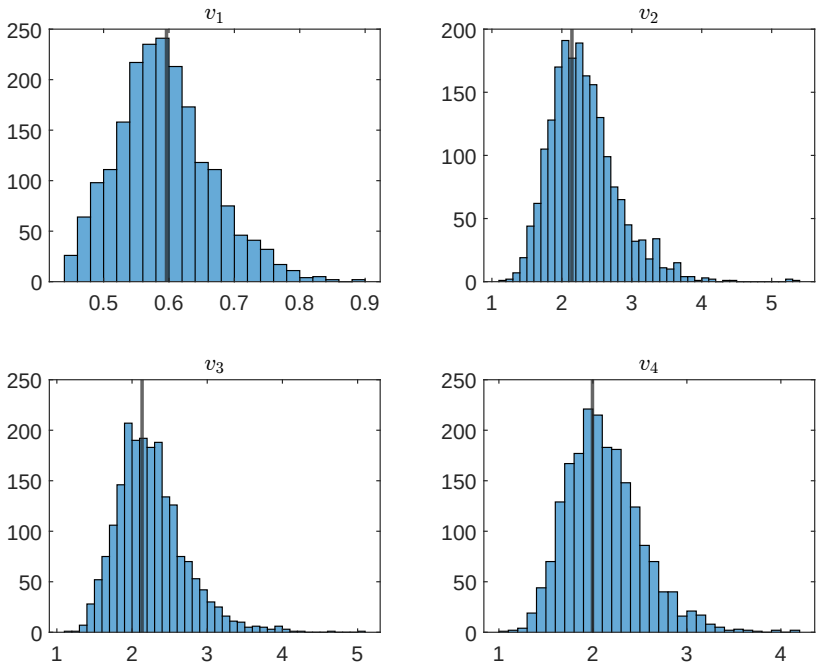

Note: Histograms of the elements of $C$ based on the posterior. Black vertical lines represent the maximum likelihood estimates.

Figure D.1 reports the posterior distribution of the elements of $C$ obtained with the simulation.

Figure D.2 reports the distribution of the degree of freedom parameters $v_1, ..., v_4$. The posterior distributions are concentrated near the maximum likelihood estimates. The prior pushes them gently towards the value of 10 but the sample information dominates and the posterior modes are very close to the maximum likelihood estimates. Almost all the probability mass lies at values below 4, implying very leptokurtic distributions. For $v_2, v_3, v_4$ significant probability mass lies at values between 1 and 2, at which the variance of the shock is infinite, and for $v_1$ virtually all probability mass lies below 1 where the variance is not defined.

Since the shocks are likely not to have a finite variance, also variance decompositions need to be taken with a grain of salt and we should expect them to be sensitive to outliers. Table D.2 reports the variance decompositions of all variables, which should be interpreted with this caveat. $u_1$ is basically equivalent to MP1. In light of this, the federal funds rate surprises are a valid instrument for the standard monetary policy shock (e.g. Kuttner (2001); Bernanke and Kuttner (2005) use this instrument). However, the most important shock is the Odyssean forward guidance shock $u_2$, which accounts for 44% of the variation of 2-year bond yields and about a half of the variation of 10-year bond yields and stock prices in

Figure D.2: Posterior distribution of $v$




Note: Histograms of $v$s based on the posterior sampler. The black vertical line represents the maximum likelihood estimate.

Table D.2: Variance decomposition

| | MP1 | ONRUN2 | ONRUN10 | SP500 |
|---|---|---|---|---|
| $u_1$ | **1.00** | **0.35** | **0.10** | **0.17** |
| | (0.00) | (0.04) | (0.03) | (0.04) |
| $u_2$ | 0.00 | **0.45** | **0.53** | **0.48** |
| | (0.00) | (0.04) | (0.10) | (0.04) |
| $u_3$ | 0.00 | 0.05 | **0.23** | 0.02 |
| | (0.00) | (0.03) | (0.10) | (0.02) |
| $u_4$ | 0.00 | **0.16** | **0.14** | **0.33** |
| | (0.00) | (0.04) | (0.03) | (0.06) |
| Total | 1.00 | 1.00 | 1.00 | 1.00 |

Note: Shares of the sample variance. Posterior standard deviations in parentheses.

the half-hour windows around FOMC announcements. The third shock that is pervasive, in the sense that it accounts for non-trivial shares of multiple variables, is the Delphic forward guidance shock $u_4$. It accounts for about 15% of the variation of Treasury yields and one third of the variation of stock prices.

Table D.3 reports the correlations between the shocks and, at the same time, illustrates the perils of applying linear statistics to non-Gaussian variables. The rank (Spearman's)

Table D.3: Rank correlations and linear correlations between the shocks

| Rank correlations | | | | | Linear correlations | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $u_1$ | $u_2$ | $u_3$ | $u_4$ | | $u_1$ | $u_2$ | $u_3$ | $u_4$ |
| $u_1$ | 1 | -0.01 | -0.01 | 0.02 | $u_1$ | 1 | -0.17 | -0.07 | -0.12 |
| $u_2$ | (0.92) | 1 | 0.02 | 0.04 | $u_2$ | (0.01) | 1 | 0.31 | 0.04 |
| $u_3$ | (0.90) | (0.77) | 1 | 0.01 | $u_3$ | (0.25) | (0.00) | 1 | 0.02 |
| $u_4$ | (0.70) | (0.52) | (0.83) | 1 | $u_4$ | (0.06) | (0.49) | (0.76) | 1 |

Note: Correlation coefficients above the diagonal, p-values in parentheses below the diagonal. Rank correlations (Spearman's correlations) in the left panel, linear correlations (Pearsons's correlations) in the right panel. The linear correlation between $u_2$ and $u_3$ drops from 0.31 to 0.06 if one omits the QE1 announcement.

correlations, reported in the left panel are all negligible. However, the linear (Pearson's) correlations, reported in the right panel, are sometimes large. Especially striking is the correlation of 0.31 between $u_2$ (forward guidance shocks) and $u_3$ (LSAP shocks). Such a high correlation between Gaussian shocks would mean that they are systematically related and hence considering their effects in isolation makes little sense. However, for non-Gaussian variables such a high linear correlations can happen by chance. In fact, in this case the linear correlation is almost entirely driven by a single observation, namely the announcement of the QE1 program in March 2009. After omitting this data point the linear correlation drops to 0.06, revealing that the shocks $u_2$ and $u_3$ are not in fact systematically linearly related.

Table D.4: Rank correlations between the *squared* shocks

| | $(u_1)^2$ | $(u_2)^2$ | $(u_3)^2$ | $(u_4)^2$ |
|---|---|---|---|---|
| $(u_1)^2$ | 1 | 0.16 | 0.12 | 0.19 |
| $(u_2)^2$ | (0.01) | 1 | 0.19 | 0.15 |
| $(u_3)^2$ | (0.07) | (0.00) | 1 | 0.23 |
| $(u_4)^2$ | (0.00) | (0.02) | (0.00) | 1 |

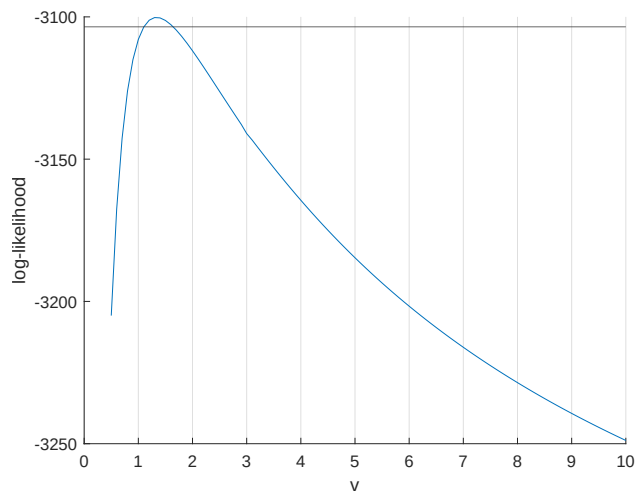
Note: Correlation coefficients above the diagonal, p-values in parentheses below the diagonal.

Table D.4 reports rank correlations between *squared* shocks, in order to understand if the shocks' absolute sizes are also independent, as assumed in model (1). It turns out that the shock sizes are not independent: in general large shocks are somewhat more likely to occur together. The rank correlations are positive and, with one exception, statistically significant at the 5% level. In light of this result, it is useful to revisit model (1) and check the robustness to relaxing the assumption of full independence (Montiel Olea et al., 2022).

# E Sensitivity of the results to the degree of non-Gaussianity

This section studies to what extent the identification weakens as we impose a higher degree of freedom parameter in the Student t distribution. It assumes a common degree of freedom parameter $\bar{v}$ for all the shocks and varies it on a grid. The results remain very similar for values of $\bar{v}$ between 1 and 12. For $\bar{v} > 12$ the identification becomes weak and the point estimates begin to change. However, these values are strongly rejected by a likelihood ratio test.

Figure E.1: Maximum log-likelihood conditional on different values of $\bar{v}$

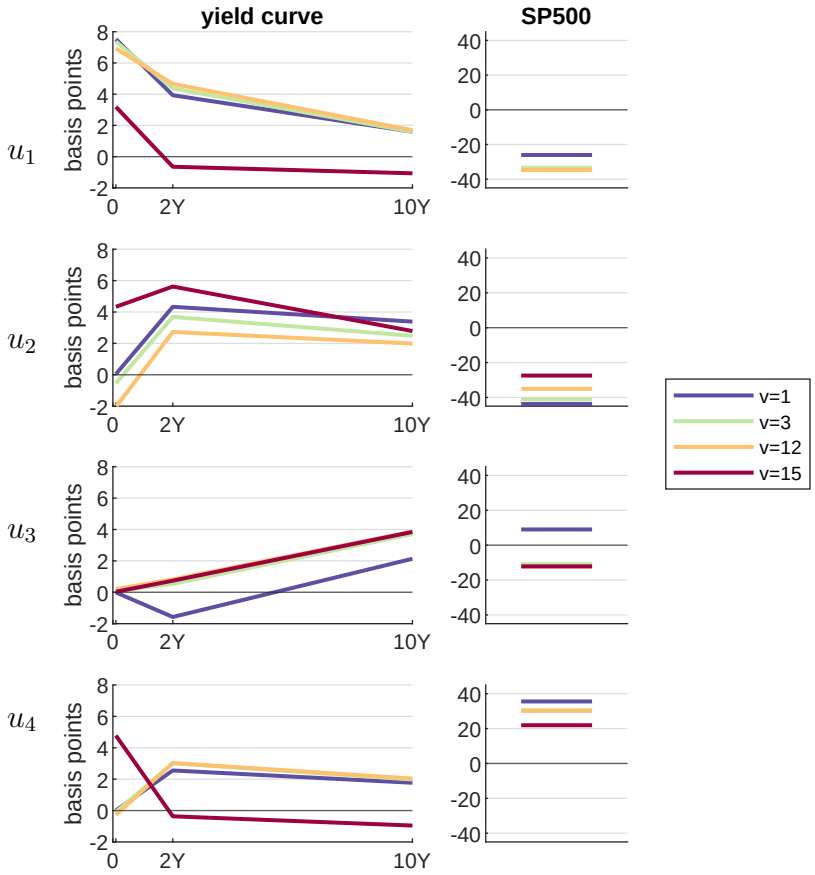



Note. The horizontal line shows the cut-off point implied by the likelihood ratio test at the 1% significance level.

More in detail, I re-estimate model (5) fixing a common value $\bar{v}$ for all shocks at a grid from 0.5 to 30. I assume that all shocks have the same, common $\bar{v}$ in order to reduce the grid size. Figure E.1 shows that the maximum attainable value of the log-likelihood decreases quickly as $\bar{v}$ deviates from the maximum at $\bar{v} = 1.33$. The figure is truncated at $\bar{v} = 10$ for readability but the log-likelihood continues to decrease also for $\bar{v} > 10$. The horizontal line at the top of the figure shows the cut-off point implied by the likelihood ratio test at the 1% significance level.

Figure E.2 shows that the effects of the four shocks are very similar for values of $\bar{v}$ between 1 and 12. Especially for the shocks $u_1$ and $u_4$ the estimates are difficult to distinguish in the figure as they lie almost on top of each other. The main visible difference is present for long-term rate shocks $u_3$: its effect on the 2-year yield is slightly negative for low $\bar{v}$ and becomes positive starting at about $\bar{v} = 3$. The point estimates change qualitatively somewhere between $\bar{v} = 12$ and $\bar{v} = 15$: shocks $u_1$ and $u_4$ become essentially fed funds rate shocks with little effect on the longer maturities, while $u_2$ becomes an almost parallel shift of the whole yield curve including the shortest maturity. However, for $v = 15$ the uncertainty (not reported here) is substantially larger and many effects are no longer statistically significant (the same is true for $\bar{v} = 12$, but not for $\bar{v} \leq 10$).

Figure E.2: The effects of standardized shocks, conditional on different values of $\bar{v}$



# F Results in subsamples

Estimation of the model on smaller sub-samples yields broadly similar results to the full sample, with two corrections to the previous messages. First, in the earlier part of the sample there is some evidence of the standard information effects associated with the movements of the current fed funds rate (as in Melosi, 2017; Nakamura and Steinsson, 2018). These standard information effects do not replace or modify the Delphic forward guidance but appear as a separate shock substituting the LSAP shock. Second, the LSAP shock $u_3$ has a significant effect on the stock prices in the later part of the sample.

Figure F.1: First vs second half of the sample

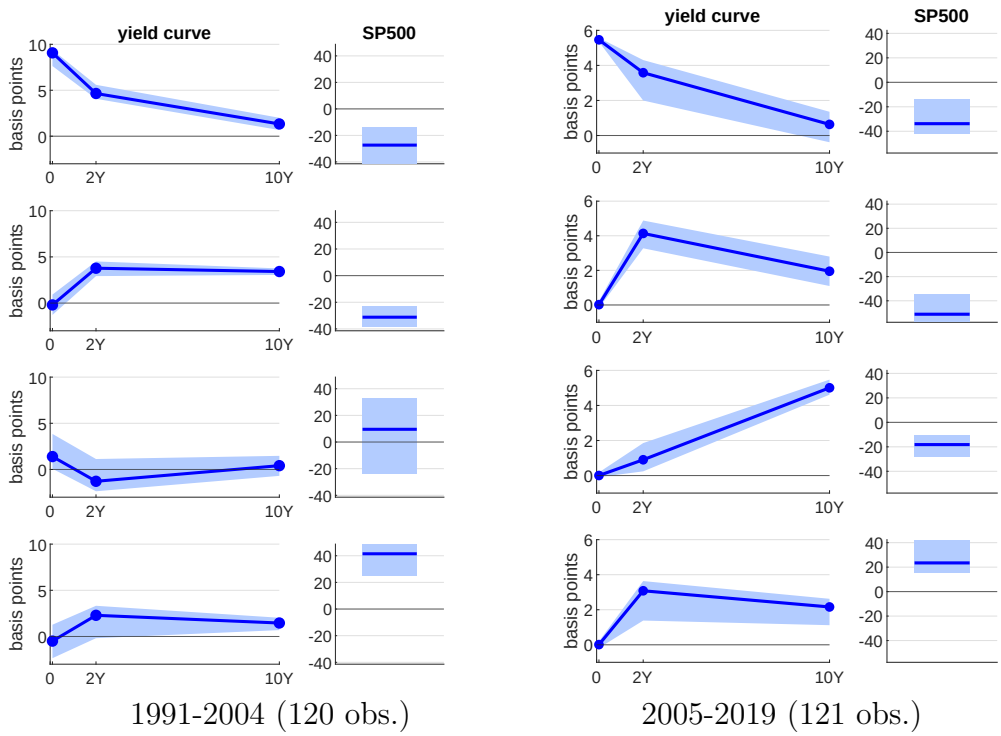
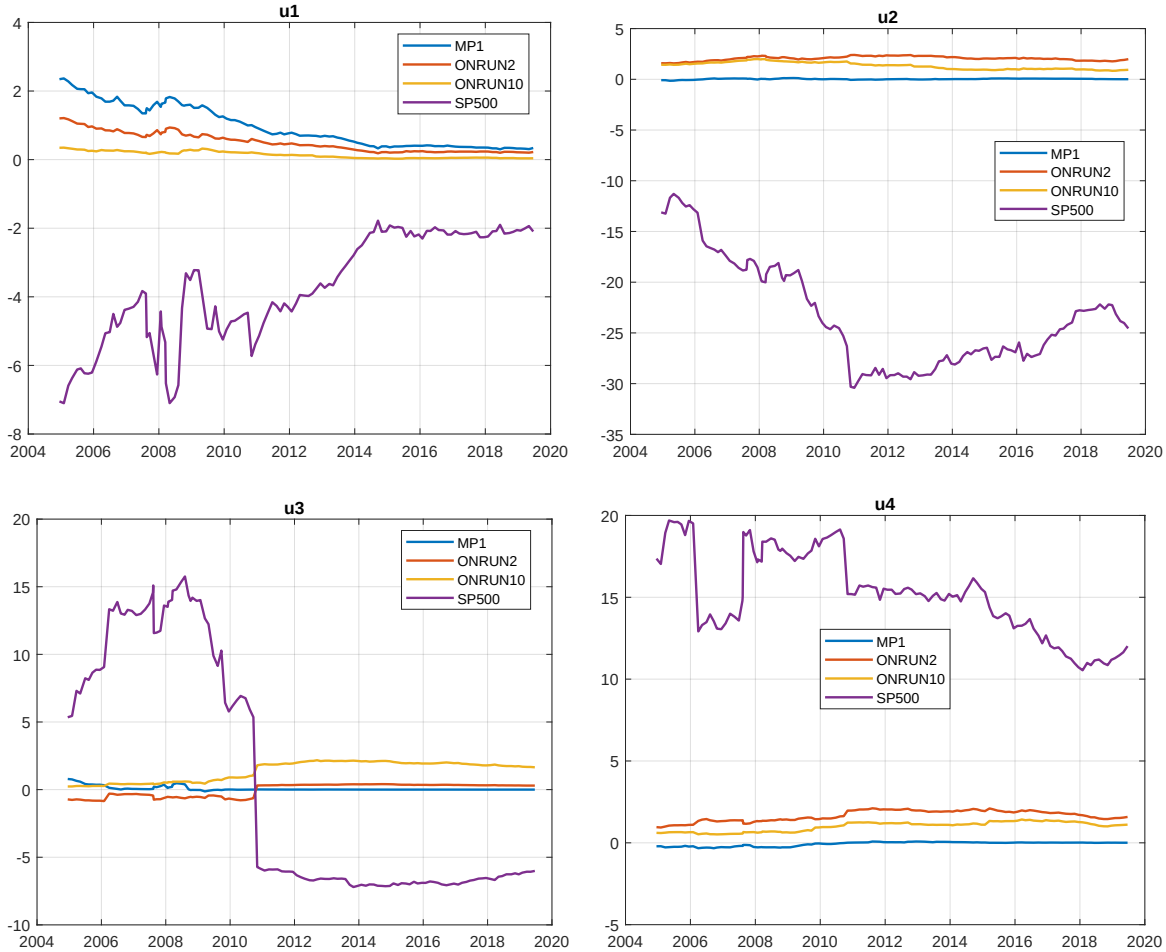



Figure F.1 reports the responses of all variables estimated in the first half of the sample (left panel) and in the second half of the sample (right panel). The error bands in these smaller samples are wider. A number of differences between the left and the right panel show up. First, the standard policy shock moves the yield curve in a similar way but is larger in the first sample (MP1 increases by 9 basis points) and smaller in the second sample (MP1 increases by less than 6 basis points). Second, in response to the Odyssean forward guidance shock $u_2$ medium and long rates move in parallel in the first sample, while the effect is hump-shaped in the second sample, with the 10-year rate moving less. Third, the LSAP shock is non-existent in the first sample. Instead, the shock $u_3$ now resembles the standard information shock associated with the fed funds rate, but is not precisely estimated. By contrast, the LSAP shock in the second sample is very pronounced and has a significant and intuitive effect on the stock prices. Finally, the Delphic forward guidance shock is broadly similar but it moves the stock prices more relatively to the interest rates in the first half of the sample.

Figure F.2: Rolling window estimates of $C$

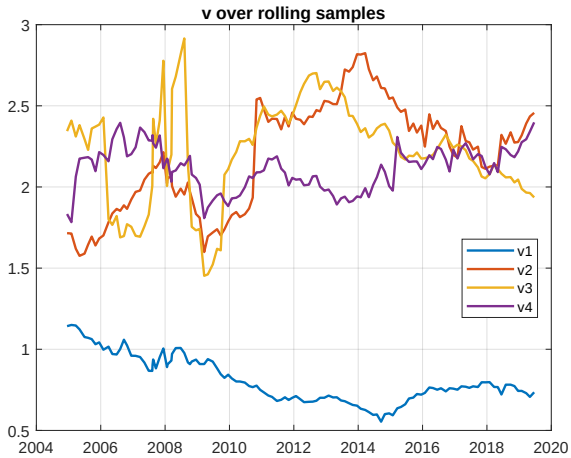



Notes. Each line plots the effect of shock $u_i$ on variable $j$, $C(i, j)$ estimated on rolling samples of 120 observations. The horizontal axis shows the last observation of the rolling sample.

Figure F.2 reports the responses of all variables estimated on rolling windows of 120 observations. Many of these models are imprecisely estimated, but the overall tendencies are clear and quite intuitive. First, the standard monetary policy shock $u_1$ becomes smaller as the windows include more observations from the ZLB period. Second, for the Odyssean forward guidance $u_2$ we can see the gradual emergence of the 'hump-shaped' yield curve response noted above. Third, the shock $u_3$ is unstable and switches from being a standard information shock in the early windows (where it is a fed funds rate hike associated with a positive stock price response) to being a contractionary LSAP shock in the later windows. The switch occurs at the point where the rolling window includes for the first time the QE1 announcement of March 18, 2009. However, the same switch occurs, only several months later, when the QE1 announcement is omitted from the sample. Finally, the Delphic forward guidance shock maintains similar features, with the elasticity of stock prices to interest rates becoming somewhat smaller in the later windows.

Figure F.3 reports the maximum likelihood estimates of $v$ obtained on the rolling windows. We can see that $v_1$ decreases as the windows include more of the ZLB period where

Figure F.3: Rolling window estimates of $v$




Notes. Each line plots the degree-of-freedom parameter $v_n$ estimated on rolling samples of 120 observations. The horizontal axis shows the last observation of the rolling sample.

standard monetary policy captured by $u_1$ becomes constrained. The other $v_n$'s fluctuate in the range between roughly 1.5 and 3.

# G Relaxing the assumption of independence

In this section I formulate and estimate an alternative model,

$$y_t = C' u_t, \qquad u_t \sim \mathcal{PDMT}(v_0, \bar{v}), \tag{G.1}$$

where $\mathcal{PDMT}(v_0, \bar{v})$ denotes the new Partially Dependent Multivariate t-distribution defined below. The PDMT distribution nests the Independent t and Multivariate t as extreme cases and spans all intermediate degrees of tail dependence between these extremes.

## G.1 The PDMT distribution

I construct the PDMT through the following steps, inspired by Jones (2002); Shaw and Lee (2008); Jiang and Ding (2016). The construction is based on the fact that a t-distributed variate can be obtained by scaling a Normal variate by an inverse square root of a Chi-squared variate divided by its degrees of freedom:

$$\text{If } z \sim \mathcal{N}(0,1), \quad q \sim \chi^2(v) \quad \text{and } t = z\sqrt{\frac{v}{q}}, \quad \text{then } t \sim \mathcal{T}(v). \tag{G.2}$$

Consequently, a vector of independent t's can be constructed as

$$\left( z_1\sqrt{\frac{v}{q_1}},\ z_2\sqrt{\frac{v}{q_2}},\ \dots \right) \tag{G.3}$$

where $z_1, z_2, ...$ are independent standard Normal variates and $q_1, q_2, ...$ are independent Chi-squared variates with $v$ degrees of freedom. The Multivariate t-distribution imposes a tight dependence on the tail behavior of all elements of the vector. A vector from the Multivariate t distribution can be constructed as

$$\left( z_1\sqrt{\frac{v}{q}},\ z_2\sqrt{\frac{v}{q}},\ \dots \right) \tag{G.4}$$

i.e. all the independent Normal variates are scaled by the same Chi-squared variate $q$. The new PDMT distribution is constructed as

$$\left( z_1\sqrt{\frac{v_0+v_1}{q_0+q_1}},\ z_2\sqrt{\frac{v_0+v_2}{q_0+q_2}},\ \dots \right) \tag{G.5}$$

where $q_0, q_1, q_2, ...$ are Chi-squared with $v_0, v_1, v_2, ...$ degrees of freedom. In the baseline case I will impose that $v_1 = v_2 = \cdots = \bar{v}$.

The $\mathcal{PDMT}(v_0, \bar{v})$ has the following attractive properties:

1. Each of its univariate marginal densities is $\mathcal{T}(v_0 + \bar{v})$. This is because the sum of a Chi-squared($v_0$) and Chi-squared($\bar{v}$) is Chi-squared($v_0 + \bar{v}$).

2. When $v_0 = 0$ it collapses to a vector of Independent t-distributions with $\bar{v}$ degrees of freedom.

3. When $\bar{v} = 0$ it collapses to a Multivariate t-distribution with $v_0$ degrees of freedom.

The disadvantage of the PDMT is that it does not have a tractable density.[1] Consequently, it needs to be studied using simulation methods.

## G.2 Estimation

I estimate model (G.1) using Bayesian methods with weakly informative priors and data augmentation.[2] I first rewrite it as

$$y_t = W^{-1\prime} Q_t^{-1/2} z_t, \quad z_t \sim \mathcal{N}(0, I_N) \tag{G.6}$$

where

$$Q_t = \text{diag}\left(\frac{q_{t0} + q_{t1}}{v_0 + \bar{v}}, \frac{q_{t0} + q_{t2}}{v_0 + \bar{v}}, \dots\right). \tag{G.7}$$

I treat the $q_{tn}$ as missing data, and specify a "prior" or a likelihood for them that is $\chi^2(v_n)$, i.e. gamma $G(v_n/2, 2)$, given by

$$p(q_{tn}) = \Gamma(v_n/2)^{-1} 2^{-v_n/2} q_{tn}^{v_n/2-1} \exp(-q_{tn}/2) \tag{G.8}$$

where $n = 0, 1, ..., N$ and, in the baseline case, $v_1 = ... = v_N = \bar{v}$. Hence, the complete data likelihood in period $t$ is

$$\begin{aligned} p(y_t, q_{1t}, ...q_{Nt} | W, v_0, ..., v_N) = |W^{-1\prime} Q_t^{-1} W^{-1}|^{-1/2} \exp\left(-\frac{1}{2} y_t' (W^{-1\prime} Q_t^{-1} W^{-1})^{-1} y_t\right) \\ \times \prod_{n=0}^{N} \Gamma(v_n/2)^{-1} 2^{-v_n/2} q_{tn}^{v_n/2-1} \exp(-q_{tn}/2) \end{aligned} \tag{G.9}$$

I specify priors for parameters $W, v_0, v_1, ....$ The prior for $W$ is flat, $p(W) \propto 1$. The prior for $v_n$ is $\mathcal{G}(\alpha_n, \beta_n) = \Gamma(\alpha_n)^{-1} \beta_n^{-\alpha_n} v_n^{\alpha_n - 1} \exp(-v_n/\beta_n)$. I use weakly informative priors $\alpha_n = 0.01$ and $\beta_n = 10$, implying the mean of 0.1 and the standard deviation of 1 (I verified that using even more diffuse priors makes no discernible difference to the results).

I conduct inference on the parameters $W, v_0, v_1, ...$ using a version of the Metropolis-Hastings algorithm. At each step of the simulation I draw new $W, v_n, q_{tn}$ for $n = 0, ..., N$ and $t = 1, ..., T$ from their respective conditional densities, conditioning on the most recent draw of the remaining quantities.

**The conditional posterior of $W$** is

$$p(W|Y, \cdot) \propto |W|^T \exp\left(-\frac{1}{2} \sum_t y_t' W Q_t W' y_t\right). \tag{G.10}$$

[1] Jones (2002) discusses a related distribution that does have a tractable density but notes that this seems to be an exception in this class of distributions.

[2] For a frequentist perspective on data augmentation see e.g. Jacquier et al. (2007).

This posterior is nonstandard. To draw from it, I draw a candidate $W^*$ from the Gaussian proposal density

$$f(W) = \mathcal{N}\left(\hat{w}, \kappa\mathcal{H}^{-1}\right) \tag{G.11}$$

where $\hat{w}$ is the mode of $p(W|Y,\cdot)$, $\mathcal{H}$ the Hessian of $\log p(W|Y,\cdot)$ and $\kappa \geq 1$ is a scalar. I derive the analytical expressions for the gradient and the Hessian of $\log p(W|Y,\cdot)$ using the methods of Magnus and Neudecker (2019).

The candidate draw is accepted with probability

$$\min\left(1, \frac{p(W^*|Y,\cdot)\, f(W)}{f(W^*)\, p(W|Y,\cdot)}\right) \tag{G.12}$$

and with the complementary probability I keep the previous draw $W$.

In the next two paragraphs I omit the subscript $t$ in $q$ to avoid clutter.

**The conditional posterior of** $q_n$, $n = 1, 2, ...N$ is

$$p(q_n|\cdot) \propto q_n^{v_n/2-1}(q_0 + q_n)^{1/2} \exp\left(-\frac{1}{2}\left(1 + \frac{u_n^2}{v_0 + v_n}\right) q_n\right) \tag{G.13}$$

This is a nonstandard density, which resembles the Gamma density except for the presence of the sum $(q_0 + q_n)$. Therefore, I draw $q_n$ from the proposal Gamma density obtained by setting $q_0$ to zero and accept the proposal draw with the probability analogous to (G.12).

**The conditional posterior of** $q_0$ is

$$p(q_0|\cdot) \propto q_0^{v_0/2-1} \prod_{n=1}^{N} (q_0 + q_n)^{1/2} \exp\left(-\frac{1}{2} q_0 \left(1 + \sum_{n=1}^{N} \frac{u_n^2}{v_0 + v_n}\right)\right) \tag{G.14}$$

Again, this density resembles the Gamma density except for the presence of the sum $(q_0 + q_n)$ and as the proposal I use the Gamma density obtained by setting $q_n = 0$ for all $n$.

**The conditional posterior of** $\bar{v}$**:** Recall that $v_1 = ... = v_N = \bar{v}$. Let $\alpha_n, \beta_n$ denote the parameters of the Gamma prior for $\bar{v}$, with the kernel $\bar{v}^{\alpha_n - 1} \exp(-\bar{v}/\beta_n)$.

$$\begin{aligned} p(\bar{v}|\cdot) \propto\ & \bar{v}^{\alpha_n - 1} \exp(-\bar{v}/\beta_n) \Gamma(\bar{v}/2)^{-TN} 2^{-TN\bar{v}/2} \prod_{n=1}^{N} \prod_{t=1}^{T} q_{tn}^{\bar{v}/2} \\ & \times (v_0 + \bar{v})^{-TN/2} \exp\left(-\frac{1}{2(v_0 + \bar{v})} \sum_{n=1}^{N} \sum_{t=1}^{T} u_{tn}^2 (q_{t0} + q_{tn})\right) \end{aligned} \tag{G.15}$$

**The conditional posterior of** $v_0$**:** Let $\alpha_0, \beta_0$ denote the parameters of the Gamma prior for $v_0$, with the kernel $\bar{v}^{\alpha_0 - 1} \exp(-\bar{v}/\beta_0)$.

$$\begin{aligned} p(v_0|\cdot) \propto\ & v_0^{\alpha_0 - 1} \exp(-v_0/\beta_0) \Gamma(v_0/2)^{-T} 2^{-Tv_0/2} \prod_{t}^{T} q_{t0}^{v_0/2} \\ & \times \prod_{n}^{N} (v_0 + v_n)^{-T/2} \exp\left(-\frac{1}{2} \sum_{n}^{N} (v_0 + v_n)^{-1} \sum_{t}^{T} u_{tn}^2 (q_{t0} + q_{tn})\right) \end{aligned} \tag{G.16}$$

The conditional posteriors of $\bar{v}$ and $v_0$ are again nonstandard densities related to the Gamma density. To draw from them I follow Jiang and Ding (2016). I first compute the conditional posterior mode and the curvature at the mode. Then I find the Gamma density with the same mode and curvature at the mode and I use that Gamma density as the proposal density.

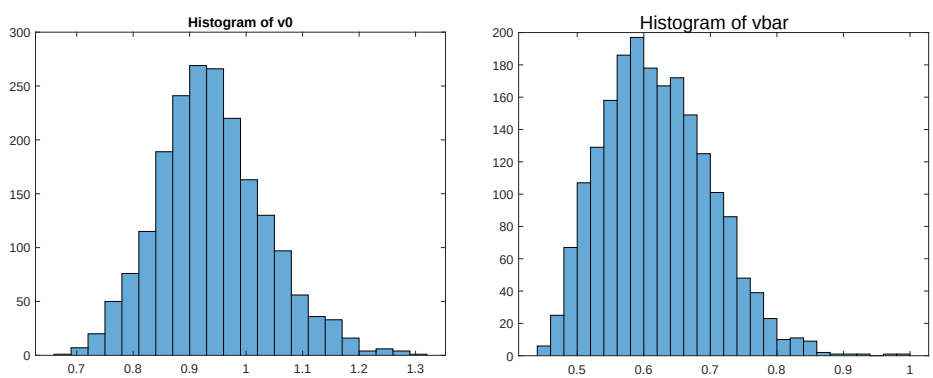
In an alternative simulation I also use the above approach of matching the mode and the curvature at the mode to drawing $q_0, q_1, ..., q_N$. In this case the simulation is substantially slower and, with a 1 million-long chain all estimation results are very similar. This suggests that the simple Gamma proposal densities for $q_n$'s are good enough.

The following results are based on a chain of 5,050,000 draws, of which I discard the first 50,000 and keep every 2500th from the rest. I confirm the stationarity of the chain with the Geweke (1992) diagnostics. While the whole inference with the Independent t model (1) in the previous section takes seconds (maximum likelihood) or minutes (simulation), generating one million draws for the PDMT takes about four hours on a standard laptop.

## G.3 Results

The PDMT model detects a nontrivial degree of tail dependence. Figure G.1 reports the posterior distributions of the degree of freedom parameters. The common degrees of freedom $v_0$ are about 50% larger than the idiosyncratic degrees of freedom $\bar{v}$ (0.9 vs 0.6).

Figure G.1: PDMT model: the posterior distributions of $v_0, \bar{v}$.

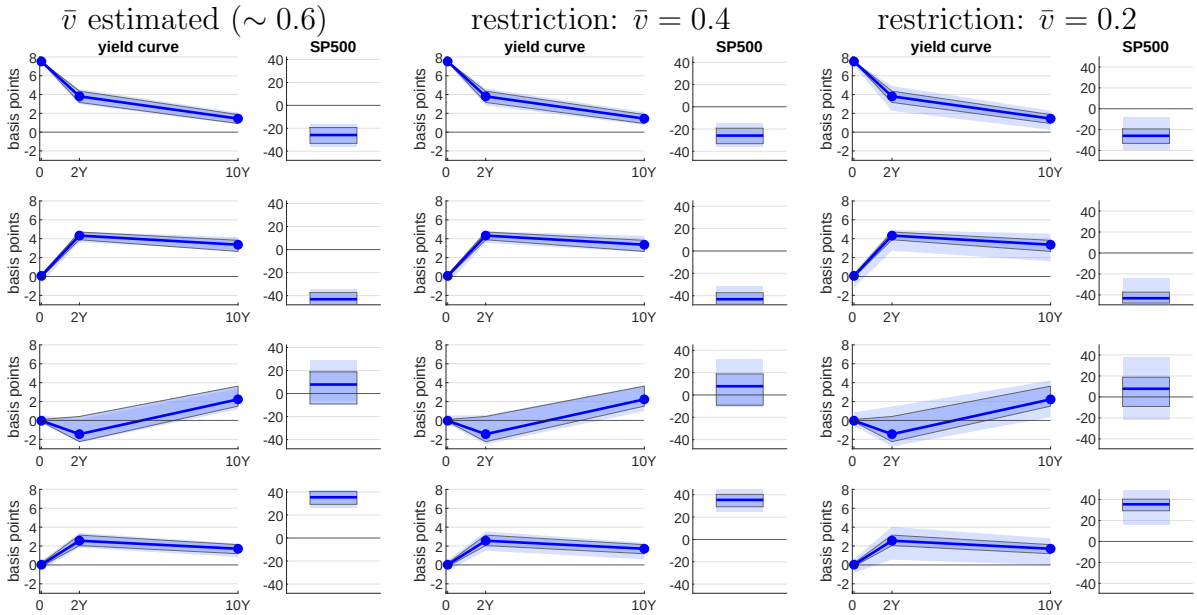



Common degrees of freedom $v_0$ Idiosyncratic degrees of freedom $\bar{v}$

The uncertainty about $C$ increases in the PDMT model, but only marginally. The first panel of Figure G.2 reports the 95% posterior bands for $C$ in the PDMT model (lighter blue), together with the 95% bands and maximum likelihood estimates in the independent Student-t model (1) for comparison (darker blue). We can see that the uncertainty bands in the PDMT model are only a little wider than the ones in the independent Student-t model. The bottom line is that the estimated shocks are the same.

To explore the sensitivity to idiosyncratic degree of freedom $\bar{v}$ I re-estimate the model imposing a restriction $\bar{v} = 0.4$ and $\bar{v} = 0.2$. As shown in panels two and three of Figure G.2, especially the last restriction widens the uncertainty bands more visibly, but the key features of the shocks are still distinguishable. Only when I try to push $\bar{v}$ even lower, the conditional posterior of $W$ becomes too flat and the algorithm runs into numerical problems. However,

Figure G.2: PDMT model: responses of the variables to standardized shocks, 95% bands.




Notes: The lighter blue areas show the 95% posterior probability bands in the PDMT model. The darker blue areas show the 95% bands in the Independent t model. The blue solid lines show the maximum likelihood estimates in the Independent t model.

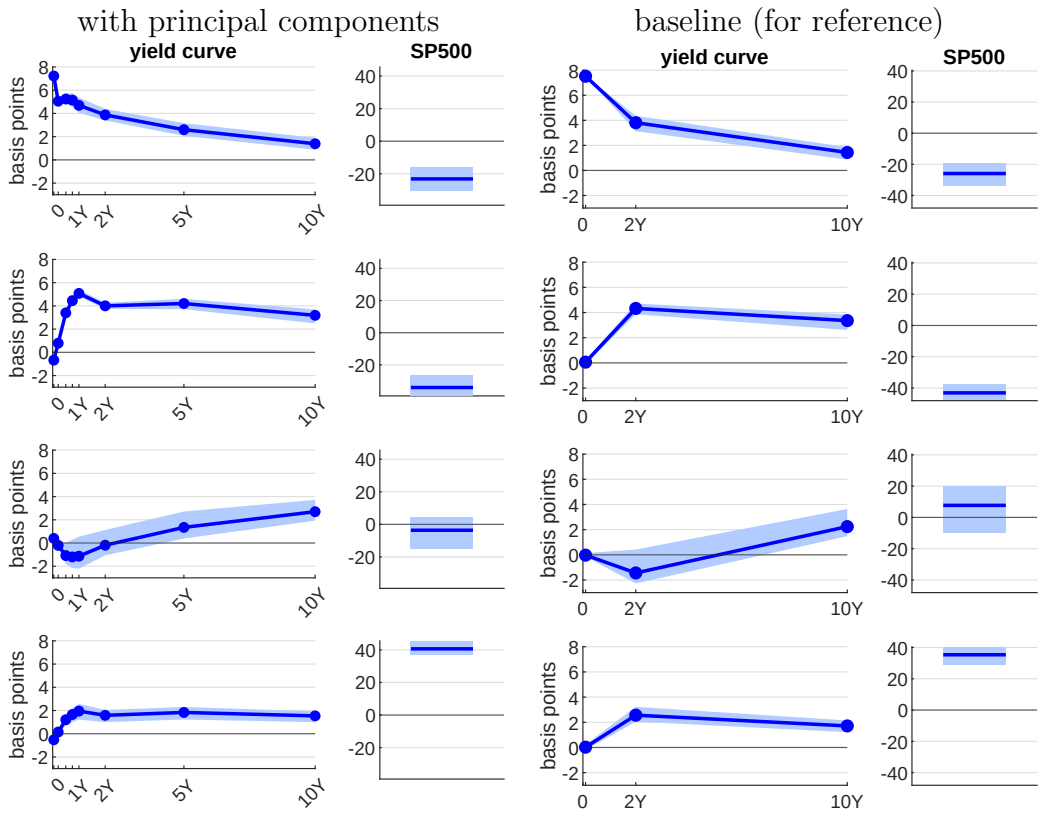
the data reject these restrictions. To study this formally, I compute the Bayes factors for comparing models with different values of $\bar{v}$ imposed. I use the bridge sampling algorithm of Meng and Wong (1996) as exposed in Müller and Watson (2020, Algorithm 4). The log Bayes factor of model $\bar{v} = 0.6$ vs $\bar{v} = 0.4$ is 3.9, indicating that 0.6 is strongly preferred to 0.4. The log Bayes factor of model $\bar{v} = 0.4$ vs $\bar{v} = 0.2$ is 26.5, indicating that 0.4 is very strongly preferred to 0.2.[3]

[3]In these model comparisons $W$ is a nuisance parameter that is common to all models and the flat prior $p(W)$ does not prevent comparing the models using Bayes factors.

# H Using more information

## H.1 Estimation with the principal components of interest rates

Figure H.1: The model with the principal components: responses of the variables to standardized shocks.




Notes: The blue areas show the 95% posterior probability bands and the solid lines with dots show the maximum likelihood estimates. The results are based on the independent t model. The bands do not take into account the uncertainty about the principal component estimation.

In this section I reestimate the baseline model replacing the three interest rate surprises MP1, ONRUN2, ONRUN10 with the first three principal components extracted from a larger set of interest rate surprises. I take eight variables from the GSS dataset: the first and third fed funds futures, the second through fourth eurodollar futures, 2-year, 5-year, and 10-year Treasury yields. In terms of the GSS identifiers, I specify

$$x = (\text{MP1, FF3, ED2, ED3, ED4, ONRUN2, ONRUN5, ONRUN10}). \tag{H.1}$$

This choice of variables follows Swanson (2021)'s choice of liquid instruments with maturities that do not overlap. I extract the first three principal components from $x$ and plug them into the model along with the SP500, i.e. I specify $y$=(PC1(x), PC2(x), PC3(x), SP500).

I estimate model (1) by maximum likelihood, obtaining four shocks and the matrix $C$ containing their effects on the three principal components and on SP500. Then I multiply the

coefficients of the principal components (i.e. the first three columns of $C$) by their loadings in the principal components analysis, thus backing out the effects of the shocks on the original GSS variables. I also repeat this for every draw from the posterior sampler. For simplicity, I abstract from the estimation uncertainty about principal components and their loadings, I treat them as known quantities. To present the responses of the variables graphically, I put the maturity on the x-axis for the yields and I put the expiration on the x-axis for the futures, so the result is not exactly a yield curve, it is just a diagram representing the movements of these variables. Figure H.1 reports the results.

We can see four shocks that are very similar as in the baseline case. Table H.2 reports their rank correlations with the baseline shocks, which range from 0.73 to 0.96. The new findings are about the intermediate maturities and futures that were missing in the baseline specification. In particular, we can see that both Odyssean and Delphic forward guidance have the strongest effects on the fourth eurodollar future, i.e. on expected 3-month LIBOR approximately one year into the future.

## H.2 Models with three shocks

In this section I estimate models with three shocks. First, I drop the SP500 surprise and limit the analysis to the three principal components of interest rate surprises, $y$=(PC1(x), PC2(x), PC3(x)). This information set is the same as in Swanson (2021). It turns out that in this case I estimate basically the same shocks as Swanson (2021). Table H.1 reports that the rank correlations between these shocks range from 0.83 to 0.95 and the linear correlations range from 0.94 to 0.97 (I normalize the lsap shock to be a tightening so for this shock the sign of the correlation is negative). Figure H.2 shows the effects of the three shocks in the first column. We can see the intuitive effects of a standard policy shock, a forward guidance shock and an asset purchase shock. It is remarkable that one can recover Swanson's shocks by maximizing the Student-t likelihood only, without imposing his bespoke factor rotations. This exercise serves as another statistical validation of Swanson's approach.

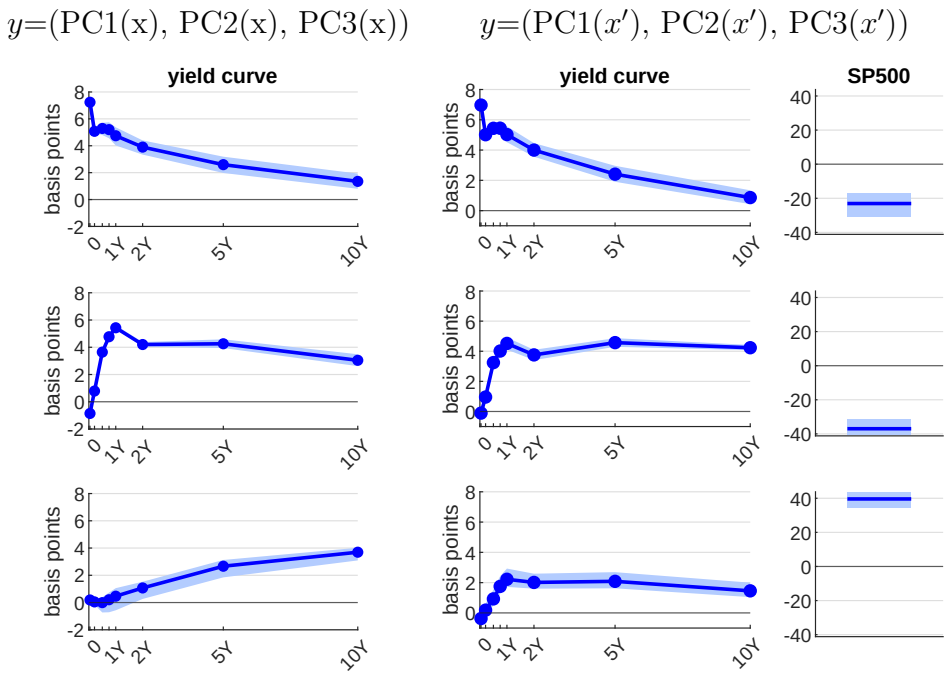
Table H.1: Pairwise rank and linear correlations for models with three shocks

| | Obs. | | $u_1$ | | $u_2$ | | $u_3$ |
|---|---|---|---|---|---|---|---|
| $y$=(PC1(x), PC2(x), PC3(x)) | | | | | | | |
| Swanson (2021) | 241 | ff: | 0.83 | fg: | 0.95 | lsap: | -0.87 |
| | | | *(0.97)* | | *(0.95)* | | *(-0.94)* |
| $y$=(PC1($x'$), PC2($x'$), PC3($x'$)) | | | | | | | |
| Baseline | 241 | $u_1$: | 0.70 | $u_2$: | 0.96 | $u_4$: | 0.97 |
| | | | *(0.94)* | | *(0.98)* | | *(0.99)* |

Note. Rank (Spearman's) correlations on top, regular font; linear (Pearson's) correlations below, in brackets, italics. 'ff', 'fg' and 'lsap' stand for fed funds, forward guidance and large scale asset purchase shocks.

In the second experiment, I include the stock price in the vector from which I extract three principal components. That is, I specify a nine-variable vector $x'$, consisting of the

Figure H.2: Models with three shocks: responses of the variables to standardized shocks.




Notes: The blue areas show the 95% posterior probability bands and the solid lines with dots show the maximum likelihood estimates. The results are based on the independent t model. The bands do not take into account the uncertainty about the principal component estimation.

previous eight variables plus SP500,

$$x' = (\text{MP1, FF3, ED2, ED3, ED4, ONRUN2, ONRUN5, ONRUN10, SP500}). \tag{H.2}$$

I extract three principal components, specify $y$=(PC1($x'$), PC2($x'$), PC3($x'$)) and estimate model (1). This time the three shocks picked up by the maximum likelihood estimation are essentially the same as the standard policy ($u_1$), Odyssean forward guidance ($u_2$) and Delphic forward guidance ($u_4$) shocks in the baseline specification. This is clear both from the impact effects of the shocks, reported in the right panel of Figure H.2 and from the high positive correlations with baseline $u_1, u_2, u_4$ reported in Table H.1.

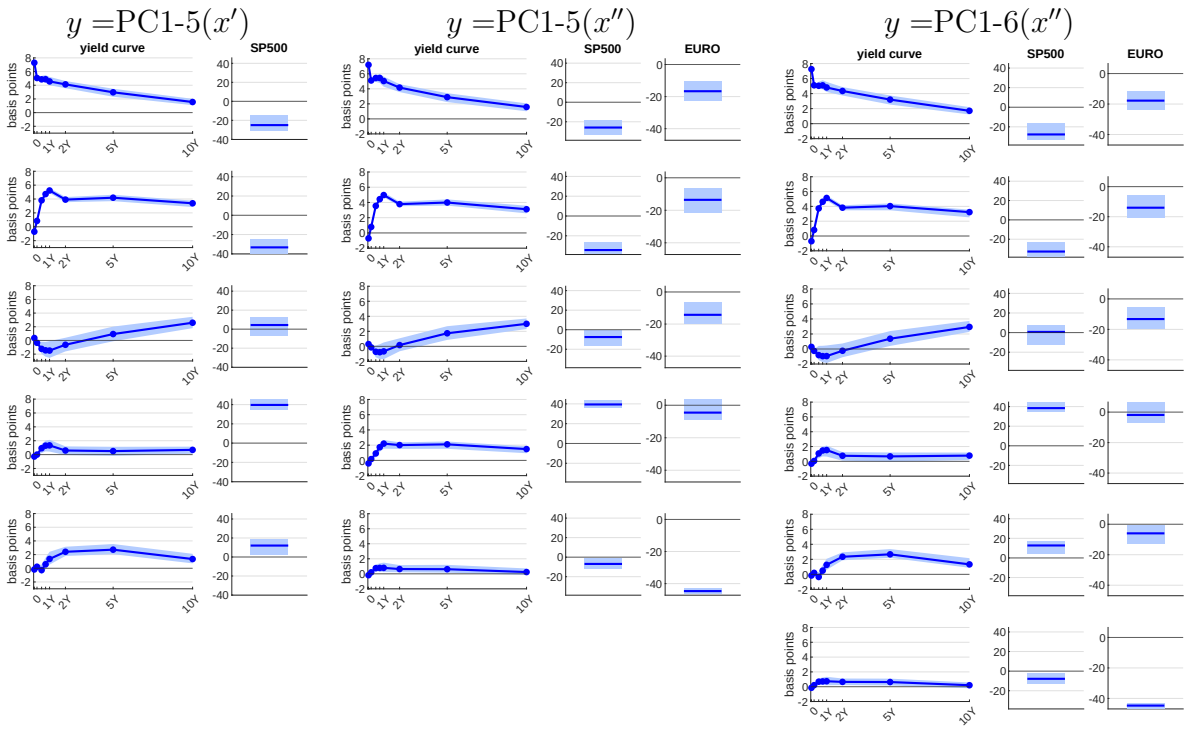
To sum up, a three shock model focused on the interest rates alone recovers the fed funds, forward guidance and LSAP shocks of Swanson (2021). A three shock model accounting for the stock prices as well recovers the fed funds, Odyssean forward guidance and Delphic forward guidance shocks.

## H.3 Searching for more shocks

In this section I estimate models with five or more shocks. These exercises yield either additional Delphic shocks differing by the stock price responsiveness, or a new shock that mainly affects the exchange rate.

First, I extract five principal components from $x'$ and specify $y$=(PC1($x'$), PC2($x'$), PC3($x'$) PC4($x'$), PC5($x'$)). See the first panel of Figure H.3. In this case the first three

Figure H.3: Models with more shocks




Notes: The blue areas show the 95% posterior probability bands and the solid lines with dots show the maximum likelihood estimates. The results are based on the independent t model. The bands do not take into account the uncertainty about the principal component estimation.

shocks remain unchanged, but instead of a single Delphic shock we now have two Delphic shocks, of which one moves the stock prices more relative to the interest rates, and another less.

Second, I specify a ten-variable vector $x''$, consisting of the previous nine variables plus the euro-dollar exchange rate,

$$x'' = (\text{MP1, FF3, ED2, ED3, ED4, ONRUN2, ONRUN5, ONRUN10, SP500, EURO}). \tag{H.3}$$

I extract five principal components from $x''$ and specify $y$=(PC1($x''$), PC2($x''$), PC3($x''$) PC4($x''$), PC5($x''$)). See the second panel of Figure H.3. In this case the first four shocks are again basically as in the baseline specification. Additionally, we can now observe the responses of the dollar. The first three shocks, standard policy, Odyssean forward guidance and LSAP shocks, have a similar effect on the dollar: it strengthens by about 15 basis points in each case. By contrast, the Delphic shock has an insignificant effect on the dollar. We also obtain a new, fifth shock which mainly affects the exchange rate, while having very small effect on the interest rates and stock prices.

In the third exercise I extract six principal components from $x''$ and include all of them in $y$. See the third panel of Figure H.3. In this case we obtain the shocks familiar from the previous two exercises: including the two Delphic shocks and the exchange rate shock, in addition to the standard policy, Odyssean forward guidance and asset purchases.

Table H.2: Pairwise rank and linear correlations with the baseline model shocks

| | Obs. | | $u_1$ | | $u_2$ | | $u_3$ | | $u_4$ |
|---|---|---|---|---|---|---|---|---|---|
| *Model in Figure H.1* | | | | | | | | | |
| PC1-3($x$),SP500 | 241 | $u_1$: | 0.87 | $u_2$: | 0.95 | $u_3$: | 0.73 | $u_4$: | 0.96 |
| | | | *(0.97)* | | *(0.97)* | | *(0.88)* | | *(0.98)* |
| *Models in Figure H.3* | | | | | | | | | |
| PC1-5($x'$) | 241 | $u_1$: | 0.89 | $u_2$: | 0.89 | $u_3$: | 0.93 | $u_4$: | 0.77 |
| | | | *(0.98)* | | *(0.94)* | | *(0.98)* | | *(0.81)* |
| | | $u_5$: | -0.04 | $u_5$: | 0.18 | $u_5$: | -0.16 | $u_5$: | 0.48 |
| | | | *(-0.24)* | | *(0.10)* | | *(-0.22)* | | *(0.61)* |
| PC1-5($x''$) | 241 | $u_1$: | 0.88 | $u_2$: | 0.90 | $u_3$: | 0.69 | $u_4$: | 0.97 |
| | | | *(0.97)* | | *(0.93)* | | *(0.88)* | | *(0.99)* |
| | | $u_5$: | -0.07 | $u_5$: | 0.17 | $u_5$: | -0.21 | $u_5$: | -0.04 |
| | | | *(-0.07)* | | *(0.10)* | | *(-0.10)* | | *(-0.05)* |
| PC1-6($x''$) | 241 | $u_1$: | 0.89 | $u_2$: | 0.87 | $u_3$: | 0.91 | $u_4$: | 0.77 |
| | | | *(0.98)* | | *(0.93)* | | *(0.98)* | | *(0.81)* |
| | | $u_5$: | -0.04 | $u_5$: | 0.17 | $u_5$: | -0.14 | $u_5$: | 0.48 |
| | | | *(-0.25)* | | *(0.10)* | | *(-0.21)* | | *(0.60)* |
| | | $u_6$: | -0.06 | $u_6$: | 0.18 | $u_6$: | -0.23 | $u_6$: | -0.05 |
| | | | *(-0.07)* | | *(0.11)* | | *(-0.11)* | | *(-0.06)* |

Note. Rank (Spearman's) correlations on top, regular font; linear (Pearson's) correlations below, in brackets, italics. The first column identifies the models by the variables included in $y$.

Table H.2 reports the rank correlations of the shocks obtained in the above exercises with the baseline shocks. In each case the first four shocks are highly correlated with the corresponding baseline shocks. The new Delphic shock is mainly correlated with the baseline Delphic shock (0.48). The new exchange rate shock is weakly negatively correlated with the LSAP shock (-0.21 to -0.23) and very little with the other shocks.

# I Further details on local projections

This section provides definitions and sources for the daily variables studied in section 6 and reports the results for additional variables.

## I.1 Daily variables: definitions and sources

The full sample is from July 1991 to June 2019. Below, the sample is only provided for the variables that are not available in the full sample.

- **Federal Funds Effective Rate** - *Source:* Fred. https://fred.stlouisfed.org Identifier: dff, after Board of Governors of the Federal Reserve System, Release H.15 Selected Interest Rates. *Units:* percent. *Transformation:* none.

- **2-year Treasury bond yield, 10-year Treasury bond yield** - Zero-coupon yield, Continuously Compounded. *Source:* https://www.federalreserve.gov/pubs/feds/2006/200628/200628abs.html Identifiers: SVENY02, SVENY10. Reference: Gürkaynak et al. (2007) *Units*: percent. *Transformation:* none.

- **S&P500** - Standard and Poor's 500 Stock Price Index *Source:* Haver. Identifier: SP500@DAILY *Units:* index. *Transformation:* 100*log.

- **High yield corporate bond OAS (US)** - ICE BofA US High Yield Index Option-Adjusted Spread (OAS). US dollar denominated below investment grade rated corporate debt publicly issued in the US domestic market. *Source:* Fred, after ICE Data Indices, LLC. Identifier: bamlh0a0hym2. *Units:* percent. *Transformation:* none. *Sample:* from 31 December 1996.

- **5-year Breakeven Inflation** - (also 2-year, 10-year, 1-year-4-years-ahead, 5-year-5-years-ahead Breakeven Inflation reported in the appendix), based on the TIPS yield curve. *Source:* https://www.federalreserve.gov/pubs/feds/2008/200805/200805abs.html Identifiers: BKEVEN05 (BKEVEN02,BKEVEN10,BKEVEN1F4,BKEVEN5F05). Reference: Gürkaynak et al. (2010) *Units*: percent. *Transformation:* none.

- **EUR per USD** - Exchange rate, extended by the BIS back to 1974. *Source:* BIS. Bilateral exchange rates, BIS WS_XRU, Identifier: D.XM.EUR.A https://data.bis.org/topics/XRU/BIS,WS_XRU,1.0/D.XM.EUR.A *Units:* Euros per one US dollar. *Transformation:* 100*log.

Variables reported in the Appendix:

- **US Short Shadow Rate (Krippner)** - References: Krippner (2013, 2015). Downloaded from https://www.ljkmfa.com/visitors/. *Units:* Percent. *Transformation:* none. *Sample:* from 3 January 1995.

- **3-month Treasury yield, gs3mo** - Market Yield on U.S. Treasury Securities at 3-Month Constant Maturity, Quoted on an Investment Basis. *Source:* Fred. https:

//fred.stlouisfed.org Identifier: dgs3mo, after Board of Governors of the Federal Reserve System, Release H.15 Selected Interest Rates. *Units:* percent. *Transformation:* none.

- **bofaml-us-aa-yld, bofaml-us-aa-oas, bofaml-us-bbb-yld, bofaml-us-bbb-oas** - ICE BofA AA US Corporate Index Effective Yield, ICE BofA AA US Corporate Index Option-Adjusted Spread, ICE BofA BBB US Corporate Index Effective Yield, ICE BofA BBB US Corporate Index Option-Adjusted Spread. *Source:* Fred, after ICE Data Indices, LLC. https://fred.stlouisfed.org Identifiers: BAMLC0A2CAAEY, BAMLC0A2CAA, BAMLC0A4CBBBEY, BAMLC0A4CBBB. *Units:* percent. *Transformation:* none. *Sample:* from 31 December 1996.

- **NFCI** - Chicago Fed National Financial Conditions Index, weekly (Fridays), *Source:* Fred, after Chicago Fed. https://fred.stlouisfed.org Identifier: NFCI. *Units:* index. *Transformation:* linearly interpolated to daily frequency.

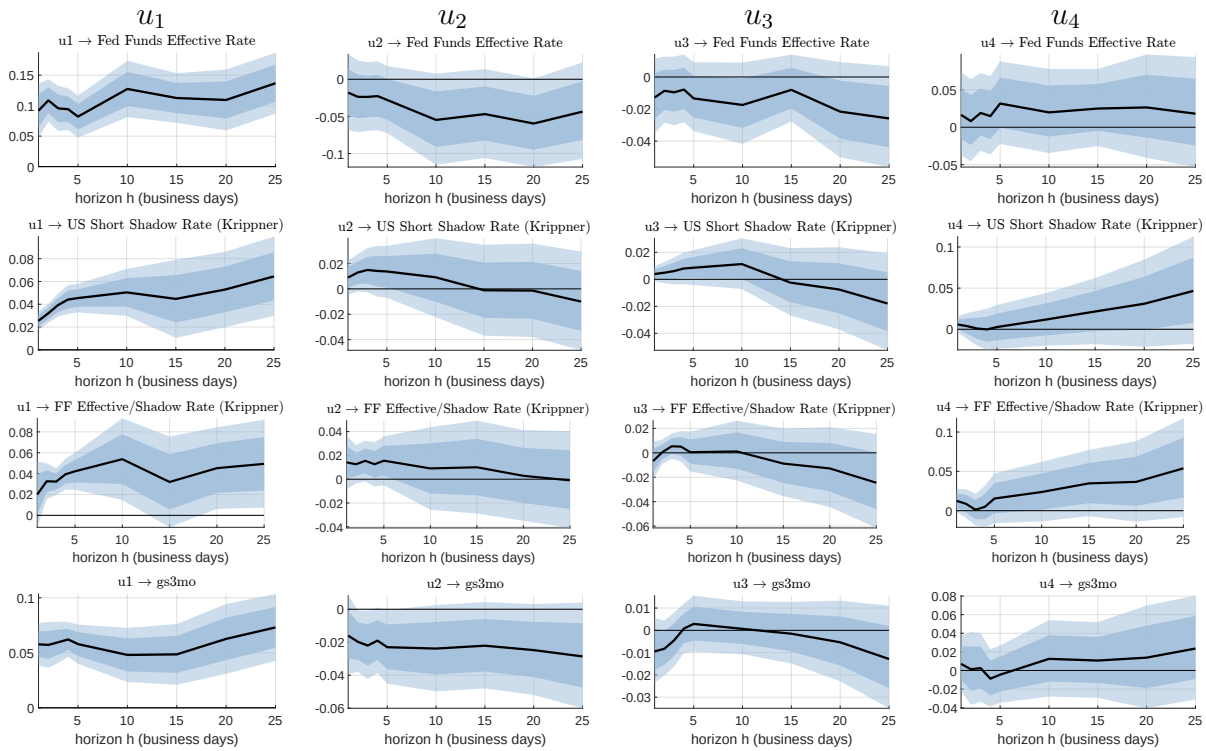
- **Broad, AFE, EME** - Nominal Broad Dollar Index, Nominal Advanced Foreign Economies Dollar Index Nominal Emerging Market Economies Dollar Index. Trade-weighted indices provided by the Federal Reserve Board. *Source:* The data from 2006 onward come from the FRB Release H10, https://www.federalreserve.gov/datadownload/Build.aspx?rel=H10. The data before 2006 come from von Beschwitz et al. (2019). *Units:* Index, per one US dollar. *Transformation:* 100*log.

## I.2 Local projection results for additional variables

Figure I.1 reports the effects of the shocks on alternative measures of short term interest rates. The first row reports the Fed Funds Effective Rate again, for reference. The second row reports the US Short Shadow Rate constructed by Krippner (Krippner, 2013, 2015). Shadow rates are designed to capture the broader monetary policy stance when the rates are constrained by the effective lower bound. We can see the intuitive result that the response of the Shadow Rate to contractionary unconventional policy shocks $u_2$ and $u_3$ becomes positive. However, it is not estimated to be very persistent. The point estimate returns to zero after about 15 business days and the results beyond the first days are not very statistically significant. Since the Shadow Rate is available only from 1995, the third row reports the results for a variable that extends the Shadow Rate using the Fed Funds Effective rate before 1995. The resulting series is not very homogeneous, as the Fed funds effective rate is significantly more volatile than the shadow rate. The responses to shocks for this variable are roughly a mix of the previous two responses. Finally the last row reports the 3-month treasury yield (gs3mo). The response to $u_1$ is dampened compared with the response of the fed funds effective rate (impact response of 6 basis points for gs3mo compared with 10 basis points for the fed funds effective rate), while the responses to the remaining shocks are similar to the fed funds effective rate.

Figure I.2 shows that corporate bond yields and spreads increase after shocks $u_1, u_2, u_3$. By contrast, after $u_4$ corporate bond spreads decline and yields do not react significantly. The first two rows show the results for the AA-rated bonds and subsequent two rows show the results for the BBB-rated bonds, which are qualitatively similar. After shocks $u_1, u_2, u_3$

Figure I.1: The effects of the shocks on daily financial variables, local projection estimates: short term interest rates



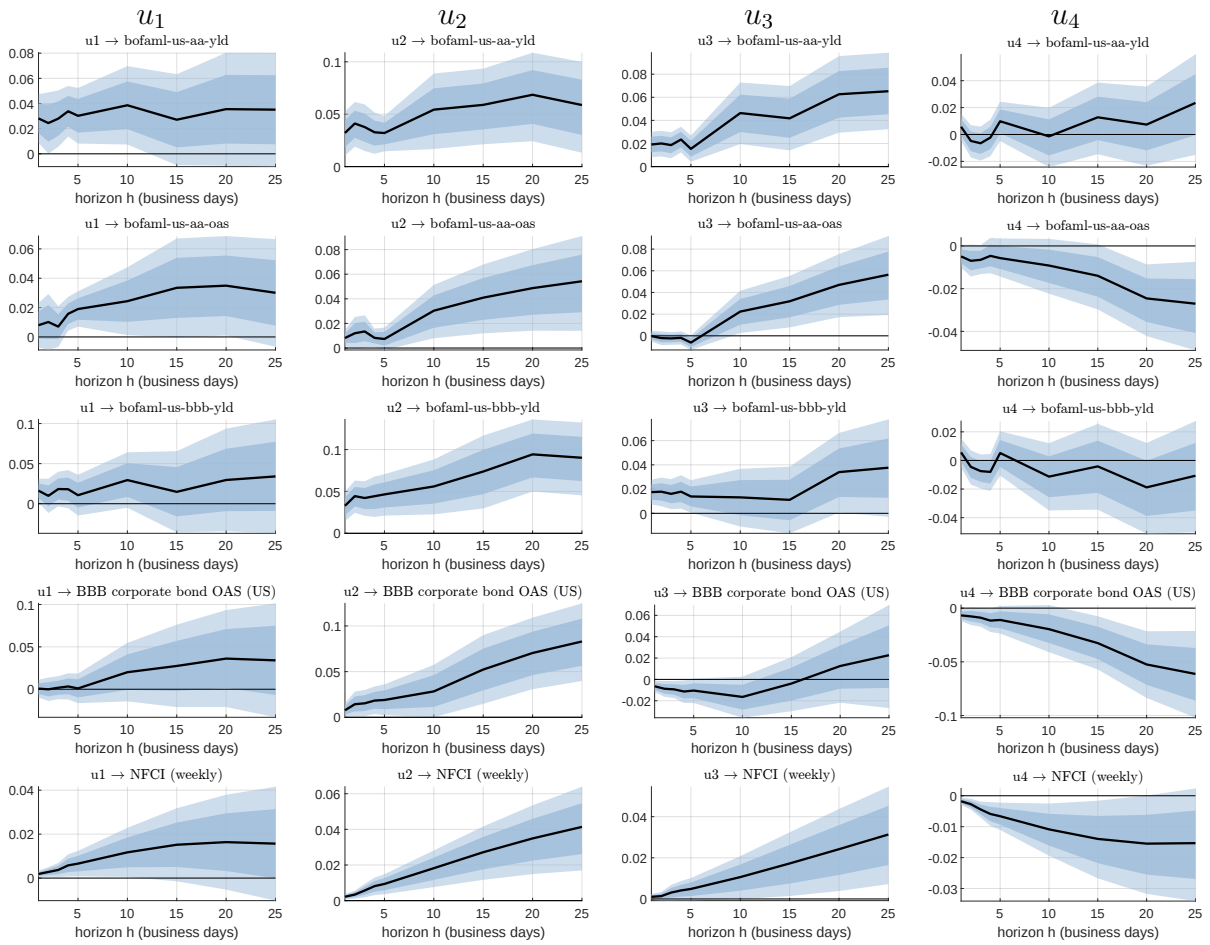

Note. See notes under Figure 5.

yields increase on impact, but the spreads do not, so the yield increases only reflect the increases in the reference Treasury yields of the same maturity. The subsequent increase in the yields is driven largely by the gradual widening of the spreads. After shock $u_4$ yields do not move significantly and spreads decline.

The last row of Figure I.2 reports the responses of the Chicago Fed's National Financial Conditions index (NFCI), which summarizes 105 financial indicators representing money markets, debt, equity markets and banking system. The story is very similar to that of the corporate bond spreads: financial conditions tighten after contractionary monetary policy shocks $u_1, u_2, u_3$ and ease after the Delphic shock $u_4$. Note that the index is linearly interpolated here from weekly to daily frequency, so it is a cruder measure of the effects, but it is reassuring that it unambiguously confirms the lessons from the corporate bond spreads.

Figure I.3 gives more information about the term structure of breakeven inflation rates. Assuming that breakeven inflation rates mainly reflect inflation expectations, the figure suggests that after $u_1$ markets expect a very gradual and persistent disinflation. We can see that the decline in the 5-year breakeven rate is larger than that in the 2-year breakeven rate and that the 1-year-4-year forward and even 5-year-5-year forward decline significantly. By contrast, after $u_2$ and $u_3$ the expected disinflation is more front-loaded, with the 2-year breakeven rate falling by most and 1-year-4-years forward mostly insignificant (in the long run markets expect inflation to rebound, as the 5-year-5-year forward increases). After $u_4$ breakeven rates increase significantly at horizons up to 5 years. The above discussion assumes

Figure I.2: The effects of the shocks on daily financial variables, local projection estimates: corporate bond yields and spreads, financial conditions index

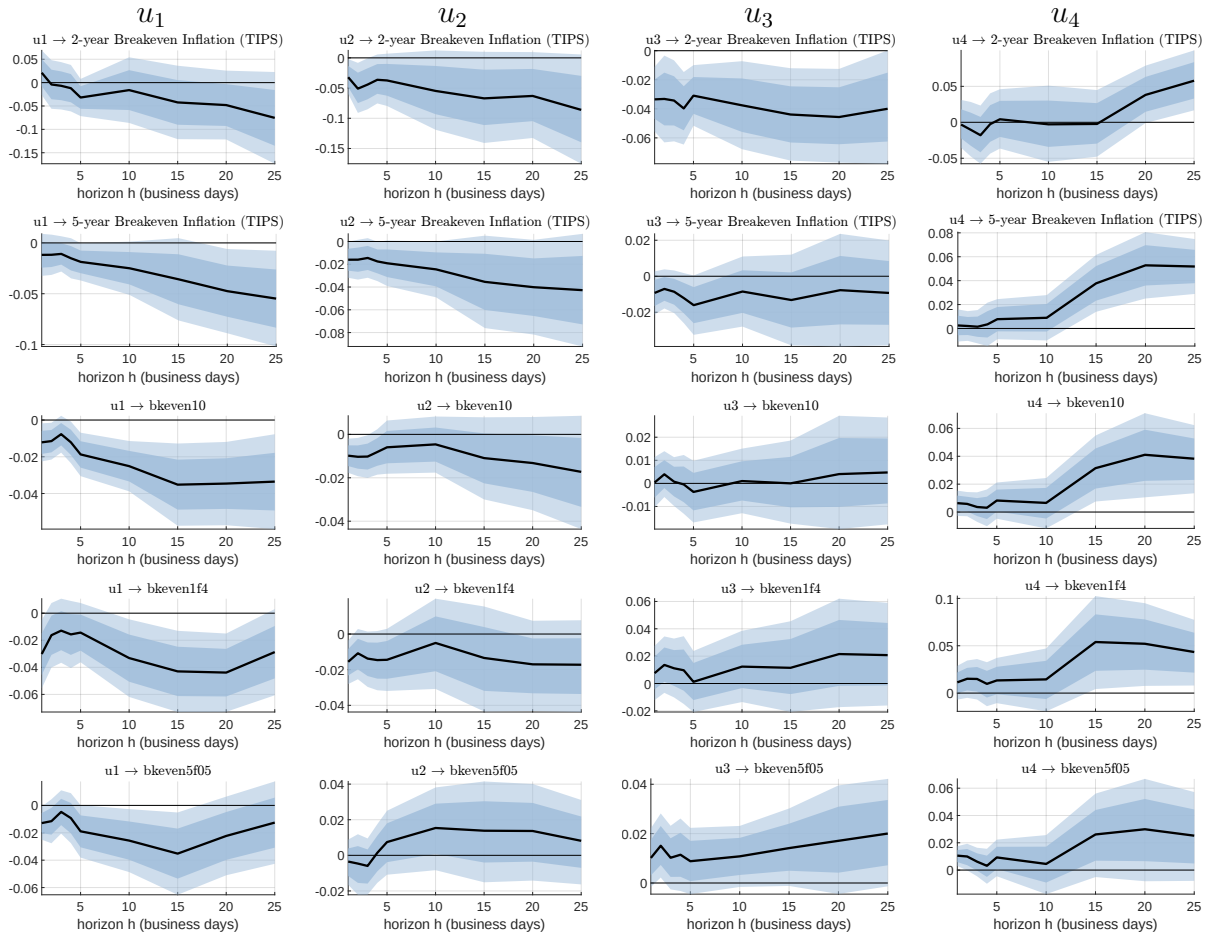



Note. See notes under Figure 5. bofaml-us-aa-yld - ICE BofA AA US Corporate Index Effective Yield; boraml-us-aa-oas - ICE BofA AA US Corporate Index Option-Adjusted Spread; bofaml-us-bbb-yld - ICE BofA BBB US Corporate Index Effective Yield; bofaml-us-bbb-oas - ICE BofA BBB US Corporate Index Option-Adjusted Spread

that breakeven rates are driven mainly by inflation expectations, the caveat should be kept in mind that they could also reflect liquidity and inflation risk premia.

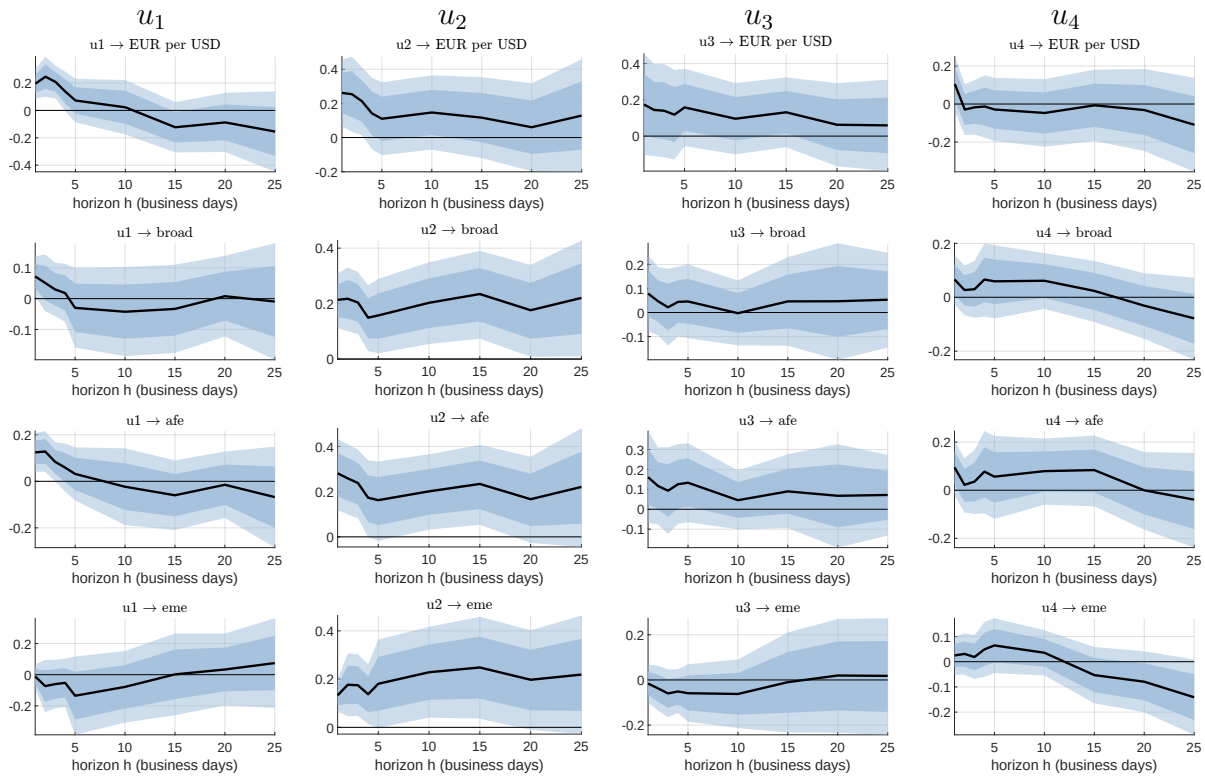
Figure I.4 reports the responses of trade-weighted nominal exchange rate indices: the broad index, the advanced foreign economies (AFE) subindex and the emerging market economies (EME) subindex. The responses to $u_2$ are the strongest and most persistent across all indices. The positive response of the broad index to $u_1$ is driven by advanced economy currencies and it is insignificant with the EME. The responses of the indexes to $u_3$ and $u_4$ are mostly insignificant. In the long run $u_4$ might even weaken the dollar against the EME currencies, consistently with the positive Delphic shock triggering an increase in the risk appetite and a capital flight *away from* safety.

Figure I.3: The effects of the shocks on daily financial variables, local projection estimates: breakeven inflation rates from TIPS




Note. See notes under Figure 5.

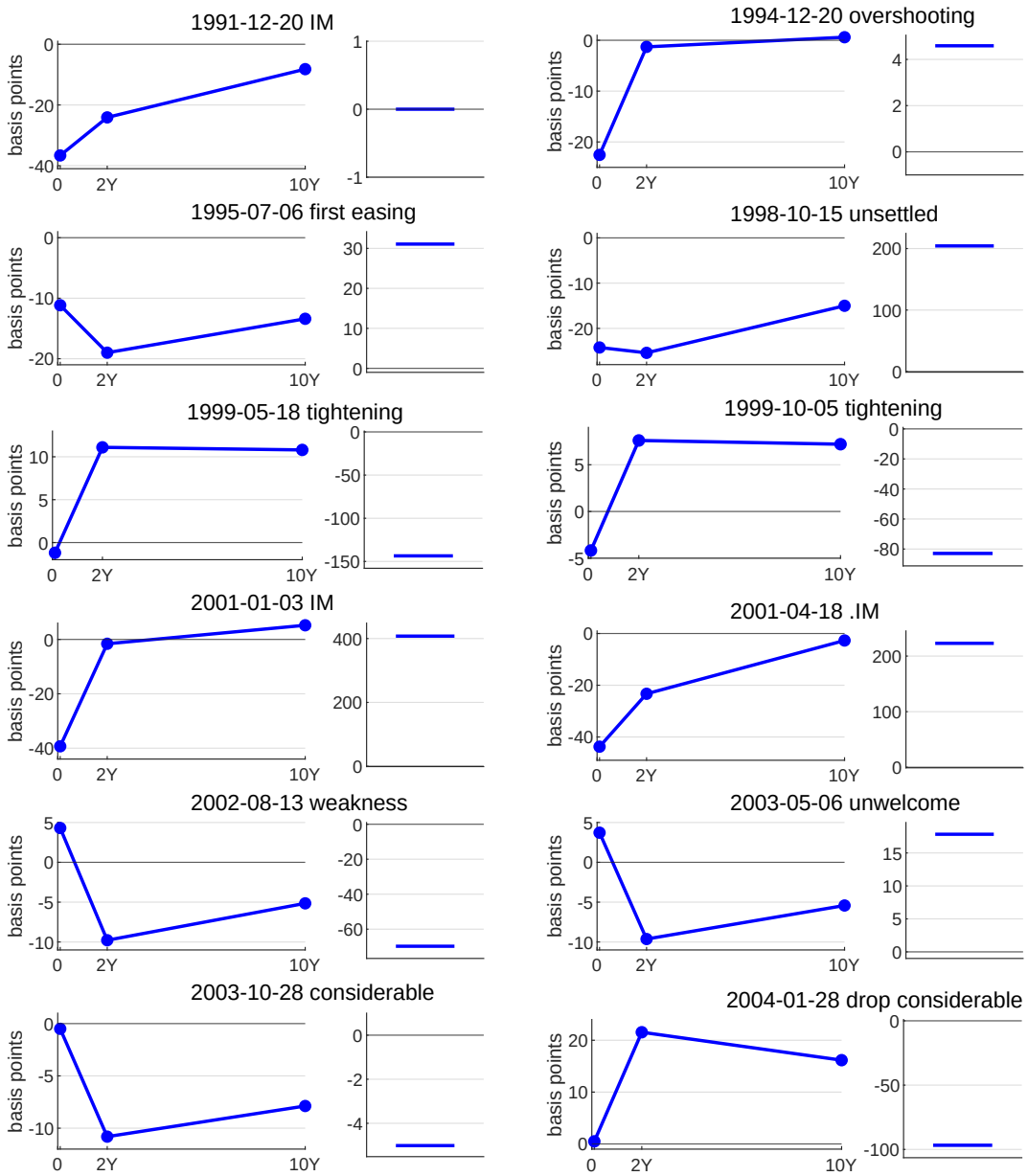
Figure I.4: The effects of the shocks on daily financial variables, local projection estimates: exchange rates




Note. See notes under Figure 5. AFE - Advanced Foreign Economies index. EME - Emerging Market Economies index.

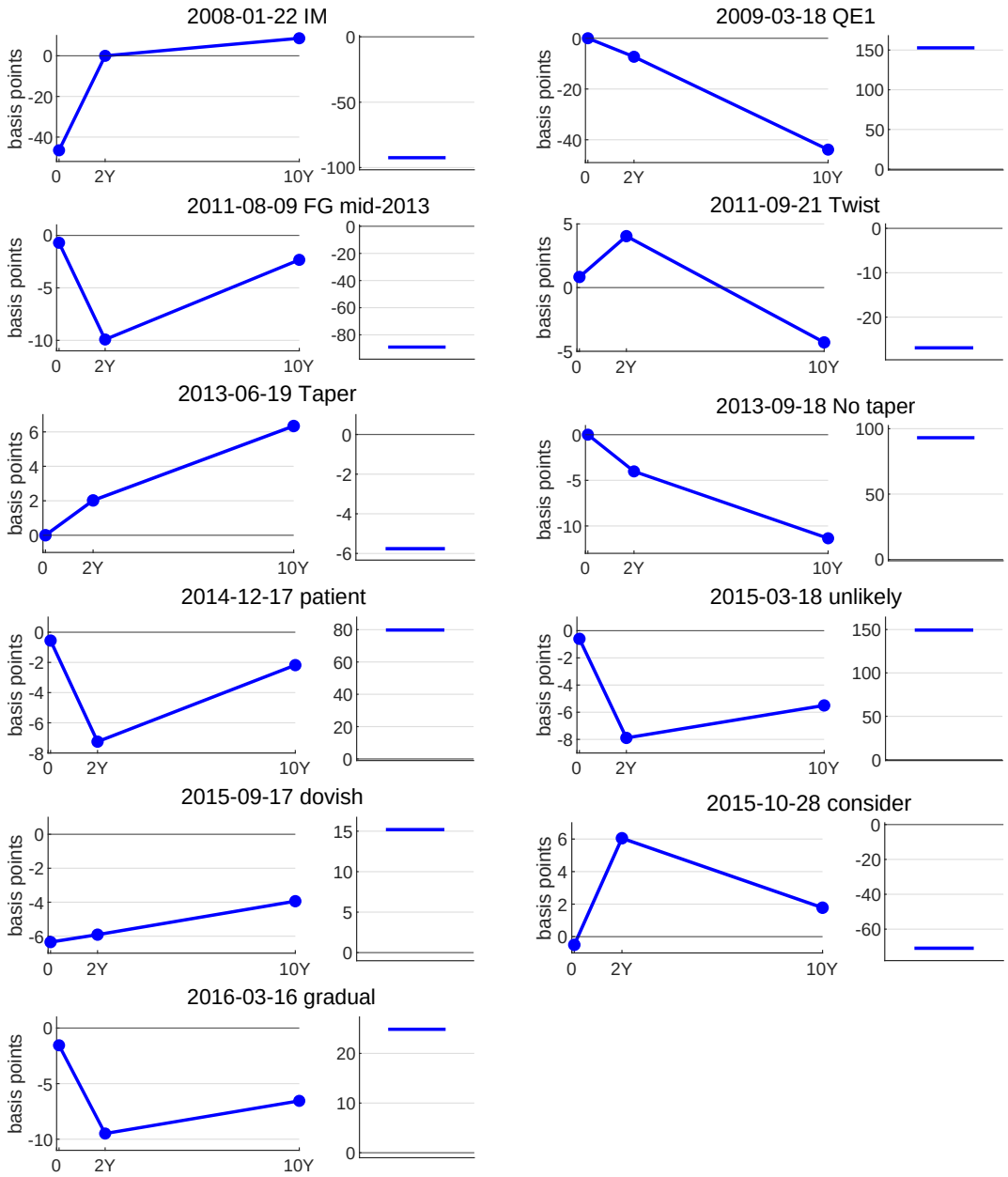
# J Additional figures

Figure J.1: The effects of selected FOMC announcements before 2008




Note. The horizontal line in the right subplots represents the change of the S&P500 stock index. IM stands for an "inter-meeting" announcement.

Figure J.2: The effects of selected FOMC announcements since 2008




Note. The horizontal line in the right subplots represents the change of the S&P500 stock index.

# References


Bauwens, Luc and Michel Lubrano (1998) "Bayesian Inference on GARCH Models Using the Gibbs Sampler," *The Econometrics Journal*, 1 (1), 23–46, 10.1111/1368-423X.11003.

Bernanke, Ben S. and Kenneth N. Kuttner (2005) "What Explains the Stock Market's Reaction to Federal Reserve Policy?" *The Journal of Finance*, 60 (3), 1221–1257, 10.1111/j.1540-6261.2005.00760.x.

Geweke, John (1992) "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments," in Berger, J.O., J.M. Bernardo, A.P. Dawid, and A.F.M. Smith eds. *Bayesian Statistics*, 169–193: Oxford University Press, Proceedings of the Fourth Valencia International Meeting on Bayesian Statistic.

Gürkaynak, Refet S., Brian Sack, and Eric Swanson (2005) "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements," *International Journal of Central Banking*, 1 (1), 55–93, https://ideas.repec.org/a/ijc/ijcjou/y2005q2a2.html.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright (2007) "The U.S. Treasury yield curve: 1961 to the present," *Journal of Monetary Economics*, 54 (8), 2291–2304, 10.1016/j.jmoneco.2007.06.029.

——— (2010) "The TIPS Yield Curve and Inflation Compensation," *American Economic Journal: Macroeconomics*, 2 (1), 70–92, 10.1257/mac.2.1.70.

Jacquier, Eric, Michael Johannes, and Nicholas Polson (2007) "MCMC maximum likelihood for latent state models," *Journal of Econometrics*, 137 (2), 615–640, 10.1016/j.jeconom.2005.11.017.

Jiang, Zhichao and Peng Ding (2016) "Robust modeling using non-elliptically contoured multivariate t distributions," *Journal of Statistical Planning and Inference*, 177, 50–63, 10.1016/j.jspi.2016.04.004.

Jones, M.C. (2002) "A dependent bivariate t distribution with marginals on different degrees of freedom," *Statistics & Probability Letters*, 56 (2), 163–170, 10.1016/S0167-7152(01)00180-8.

Khatri, C. G. and C. Radhakrishna Rao (1968) "Solutions to some functional equations and their applications to characterization of probability distributions," *Sankhya*, 30, 167–180.

Krippner, Leo (2013) "Measuring the stance of monetary policy in zero lower bound environments," *Economics Letters*, 118 (1), 135–138, 10.1016/j.econlet.2012.10.011.

——— (2015) *Zero Lower Bound Term Structure Modeling: A Practitioner's Guide*: Palgrave-Macmillan, 10.1057/9781137401823.

Kuttner, Kenneth N. (2001) "Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market," *Journal of Monetary Economics*, 47 (3), 523–544, 10.1016/S0304-3932(01)00055-1.

Magnus, Jan R. and Heinz Neudecker (2019) *Matrix Differential Calculus with Applications in Statistics and Econometrics*: John Wiley, 3rd edition, 10.1002/9781119541219.

Melosi, Leonardo (2017) "Signalling Effects of Monetary Policy," *The Review of Economic Studies*, 84 (2), 853–884, 10.1093/restud/rdw050.

Meng, Xiao-Li and Wing Hung Wong (1996) "Simulating ratios of normalizing constants via a simple identity: a theoretical exploration," *Statistica Sinica*, 6 (4), 831–860, http://www.jstor.org/stable/24306045.

Montiel Olea, José Luis, Mikkel Plagborg-Møller, and Eric Qian (2022) "SVAR Identification From Higher Moments: Has the Simultaneous Causality Problem Been Solved?" *AEA Papers and Proceedings*, 112, 481–85, 10.1257/pandp.20221047.

Müller, Ulrich K. and Mark W. Watson (2020) "Low-Frequency Analysis of Economic Time Series," Working Papers 2020-13, Princeton University. Economics Department. http://www.princeton.edu/~umueller/.

Nakamura, Emi and Jón Steinsson (2018) "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect," *The Quarterly Journal of Economics*, 133 (3), 1283–1330, 10.1093/qje/qjy004.

Rubio-Ramírez, Juan F., Daniel F. Waggoner, and Tao Zha (2010) "Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference," *Review of Economic Studies*, 77 (2), 665–696.

Shaw, William T. and K.T.A. Lee (2008) "Bivariate Student t distributions with variable marginal degrees of freedom and independence," *Journal of Multivariate Analysis*, 99 (6), 1276 – 1287, 10.1016/j.jmva.2007.08.006.

Swanson, Eric T. (2021) "Measuring the Effects of Federal Reserve Forward Guidance and Asset Purchases on Financial Markets," *Journal of Monetary Economics*, 118, 32–53, 10.1016/j.jmoneco.2020.09.003.

von Beschwitz, Bastian, Christopher G. Collins, and Deepa D. Datta (2019) "Revisions to the Federal Reserve Dollar Indexes," FEDS Notes January 2019, Board of Governors of the Federal Reserve System, 10.17016/2573-2129.48.

Waggoner, Daniel F. and Tao Zha (2003) "Likelihood preserving normalization in multiple equation models," *Journal of Econometrics*, 114 (2), 329–347, 10.1016/S0304-4076(03)00087-3.